An Atomistic Picture of the Active Interface in Dye Sensitized Solar Cells

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde (Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Florian Schiffmann

aus

Deutschland

Promotionskomitee

Prof. Dr. Jürg Hutter (Vorsitz)

Dr. Joost VandeVondele (Leitung der Dissertation)

Prof. Dr. Peter Hamm

Zürich, 2010

Abstract

The objective of this work was to contribute to a detailed understanding of the atomistic structure and the processes of the active interface in dye sensitized solar cells (DSSC) using computational chemistry methods. This study deals with a prototypical setup used in high performance DSSC, consisting of anatase nanocrystals, the cis-bis(4,4'-dicarboxy-2,2'-bipyridine) dithiocyanato ruthenium(II) dye (N3) and the iodide/triodide redox couple in acetonitrile as liquid electrolyte. Using force field based molecular dynamics (FF-MD), the structure of the anatase acetonitrile interface is studied. A strong interaction of the first layer of acetonitrile with the (101) surface of anatase is found which passivates the surface. In combination with the strong dipole of acetonitrile, the ordering of the first layer propagates up to 12Å into the bulk solvent and leads to a reduced self diffusion constant in this region. From density functional theory (DFT) calculations, a pH dependent equilibrium for the preferred binding mode of the dye on the anatase surface is found, and could be verified by attenuated total reflection (ATR)-IR experiments. Furthermore, a chain-like self-assembly via hydrogen bonds between neighboring dyes is possible for the determined binding mode.
Next to these structural insights, the regeneration mechanism of the N3 dye has been determined by ab initio molecular dynamics. This mechanism is based on an interaction of the thiocyanate ligand with iodide or diiodide, forming weakly bound intermediate complexes. The predicted formation of a N3-diiodide complex on the anatase surface has been verified afterwards using ATR-IR experiments. Furthermore, the concentration profile of the redox couple close to the anatase surface has been computed. An inhomogeneous distribution of iodide was found with a maximum concentration in the region of the thiocyanate ligand, explaining the fast regeneration in real devices. Finally, by systematic studies of system size, solvation effects and density functionals the minimal requirements for unbiased calculations of electronic and excited state properties have been determined. This setup has then been used to compute the excitation spectrum of the N3-dye on the surface and to estimate a lower bound for the electron injection rate of the N3-dye.

Zusammenfassung

Ziel dieser Arbeit war es, mit Hilfe von Methoden der Computerchemie zu einem detaillierten Bild der atomaren Struktur und den Prozessen im Bereich der aktiven Grenzfläche in Farbstoff-Solarzellen (DSSC) beizutragen. Die Untersuchungen in dieser Arbeit beschränken sich auf die Komponenten einer prototypischen Hochleistungszelle, bestehend aus Anatas Nanokristallen, dem cis-bis(4,4'-dicarboxy-2,2'-bipyridine) dithiocyanato ruthenium(II) Farbstoff (N3) und dem Iodid/Triiodid Redoxpaar in Acetonitril als Flüssig-Elektrolyt. Die Struktur des Anatas-Acetonitril an der Grenzfläche wurde mittels kraftfeldbasierter Moleküldynamik (FF-MD) untersucht. Es wurde eine starke Wechselwirkung zwischen der ersten Schicht Acetonitril und der Anatase(101) Oberfläche gefunden, die zu einer Passivierung der Oberfläche führte. Durch den starken Dipol von Acetonitril setzt sich die Struktur der ersten Schicht bis 12Å in das Lösungsmittel fort und resultiert in diesem Bereich in einer stark verringerten Selbstdiffusionskonstante. Dichtefunktional Theorie Berechnungen zur Identifizierung der Bindungsstruktur des Farbstoffes an der Oberfläche zeigten ein pH-abhängiges Gleichgewicht zweier Strukturen, das durch attenuated total reflection (ATR) IR Experimente bestätigt werden konnte. Desweiteren konnte gezeigt werden, dass eine kettenförmige Selbstassemblierung durch Wasserstoffbrücken benachbarter Moleküle in dieser Struktur möglich ist.
Neben strukturellen Eigenschaften konnte durch ab initio Moleküldynamiksimulationen, der Mechanismus, über den der N3-Farbstoff regeneriert wird, aufgeklärt werden. Dieser basiert auf einer Wechselwirkung zwischen dem Thiocyanat-Liganden und Iodid oder Diiodide, die zur Bildung eines schwachgebundenen intermediären Komplexes führt. Mittels ATR-IR Experimenten konnte der vorhergesagte Diiodid-N3 Komplex auch experimentell nachgewiesen werden. Ferner wurde das Konzentrationsprofil des Redoxpaares an der Grenzfläche berechnet. Eine stark inhomogene Verteilung wurde gefunden, die ein Maximum im selben Abstand wie der SCN-Ligand aufweist, was die schnelle Regeneration des Farbstoffes in echten Zellen erklärt. Schliesslich wurden durch eine systematische Analyse von Grösseneffekten, Solvatisierungseffekten und Dichtefunktionalen die minimalen Anforderungen zur Berechnung der Elektronenstruktur und Eigenschaften der angeregten Zustände bestimmt. Dies wurde im folgenden benutzt um das Anregungsspektrum und ein oberes Limit für die Elektroneninjektionszrate des N3-Farbstoffes auf der Oberfläche zu bestimmen.

Acknowledgement

I would like to thank Professor Jürg Hutter for giving me the opportunity to do my PhD in his group. Whenever there has been a problem, his door was open.

It is difficult to overstate my gratitude to my supervisor Dr. Joost VandeVondele. Not only in the very beginning, when he taught me the fundamentals of molecular dynamics, but throughout the whole period, he was always patient, provided me good advice and his support in understanding difficult problems. I really enjoyed the collaboration, and I know without him, this work would have been much harder.

I am indebted to Dr. Atsushi Urakawa, Dr. Ronny Wirz and Professor Alfons Baiker, who were responsible for the IR-experiments. They never got tired discussing their results with me, which finally allowed for an interpretation being an important part of my work.

I would like to thank my colleague and friend Manuel Guidon for all support and new ideas and insight concerning computers. I will miss the coffee breaks with him and Joost, as they have always been either entertaining or scientifically interesting, but most of the time both.

Furthermore, I want to thank Dr. Ari Seitsonen, Dr. Urban Borstnik and the rest of the working group for the being always cooperative and the great atmosphere in the group.

Finally, I want to express my gratitude to all other developers of CP2K, as in my opinion it is a great program they have written.

Contents

Chapter 1

Introduction

Dye sensitized solar cells (DSSC), are a relatively new class of solar cells, reaching up to 11% efficiency under laboratory conditions. Contrary to conventional cells, the photosensitive part of these cells is not the semiconducting material but a dye adsorbed on it. The production of the photosensitive semiconductor is the most expensive and energy-intensive part in the assembly of conventional devices. The separation of the semiconducting and the photosensitive part therefore reduces the demands on the semiconductor and cheaper materials can be used. Hence, the energy invested for production is thus returned much faster in DSSC.
Next to all these desirable features, the composition of DSSC is more complex than of conventional devices. As the sensitizing dye injects an electron into the semiconductor after photoexcitation, an electrolyte is needed to transport the charge from the counterelectrode back to the dye. Therefore, the efficiency of these devices is not only determined by the semiconductor and the dye but by all possible interactions arising from the additional substances. For a systematic improvement of the DSSC efficiency an as detailed as possible picture of the photoactive interface is crucial.
In the past 20 years, various experimental studies studies have been performed to improve the understanding of the processes and the influence of the various interactions in these devices. Several important features could be identified this way, but a detailed atomistic picture of the structure and the interplay between the different compounds is still missing. One of the aims of this work is to contribute to a better understanding of the structure and the processes in DSSC using computational chemistry methods. The development of a more realistic model requires an adequate description of the semiconductor surface, as well as the sensitizers and the electrolytes interaction with it. Therefore, the minimal requirements in terms of the level of theory and system size have been carefully determined to obtain unbiased results. In this work detailed studies on the various interfaces are performed, finally leading to a setup resembling the most important features of a real device.

This thesis is divided into eight chapters. After this introduction, the sec-

ond chapter will provide a more detailed introduction to DSSC. It will start with a short history of DSSC research and afterwards discuss the different components and materials used in today's devices. Finally, an overview of the most important questions and challenges concerning DSSC will be given. In the next chapter the basic electronic structure methods used in this work will be described. Furthermore, new methods developed or implemented in the CP2K program package as part of this work will be presented, and some details about the implementation, validation and performance of these methods will be given.
The following four chapters will present computed properties of the active interface, starting in chapter four with an analysis of the structure and dynamics of the acetonitrile-liquid interface, which plays an important role in DSSC using a liquid electrolyte. As part of this, a parameterization of a classical force field for the acetonitrile anatase interaction will be derived. Chapter five focuses on the binding mode of N3-like[1] dyes on the anatase (101) surface and the meaning of its orientation for the packing of these dyes in DSSC. In chapter six, the regeneration mechanism for the N3-dye by the iodide triiodide redox couple will be studied, and a mechanism avoiding biradicalic diiodide reactions will be obtained. Furthermore, the distribution of the redox couple close to the anatase surface will be studied and an explanation for its superior performance will be given. The next chapter focuses on the modeling of electronic and excited state properties of the active interface. A detailed analysis of the influence of the level of theory and system setup will be presented. The quality of the proposed setup is then tested by the computation of excitation spectra of the dye on the surface and electron injection properties of the dye. In the final chapter the results will be summarized and a short discussion of their meaning for the development of new devices will be discussed.

Chapter 2

Dye Sensitized Solar Cells

The working principle of dye sensitized solar cells (DSSC) originates from the late 1960s[2]. The first cells able to produce an electric current were based on single crystal wide band gap semiconductors and chlorophyll and its derivatives as sensitizing dyes. With an incident energy conversion efficiency (η) of less than 0.5%, these cells have been a proof of concept, but were far away from real applications. Until 1991 the efficiency of the cells was improving very slowly reaching an efficiency of 2.5% at maximum[3]. In 1991 the group of Grätzel presented a DSSC based on sintered TiO_2 nanoparticles sensitized by a ruthenium dye with an efficiency exceeding 7%, respectively 10% in 1993[4, 1]. Proving that these cells can reach efficiencies required for practical applications, DSSC became an important research topic.
This chapter will discuss the technology of DSSC's. After giving a short introduction on the general physical processes in DSSC's, it follows a more detailed description of the different materials and their properties used in modern devices. Finally the current status of laboratory and commercial cells, the problems, open questions, and challenges for the further development of DSSC's will be discussed.

2.1 Operating Principle

The DSSC devices are a complex composition of various compounds. The simplest working active interface consists of an electrode (transparent conducting glass, conducting polymers, ...) coated with a porous nanocrystalline wide bandgap semiconductor like TiO_2 (mainly anatase exposing the (101) surface) or ZnO on which a suitable dye is chemisorbed. The counter electrode is usually the same material as the active electrode, but coated with a redox catalyst (e.g. Pt). The charge transport from the counter electrode in liquid electrolyte DSSC happens via a redox mediator like the I^-/I_3^- couple or either a hole conducting polymer or a p-type semiconductor in solid state devices.
In Figure 2.1 a schematic representation of the working principle is shown. Initially, an electron of the sensitizing dye is excited by an incident photon

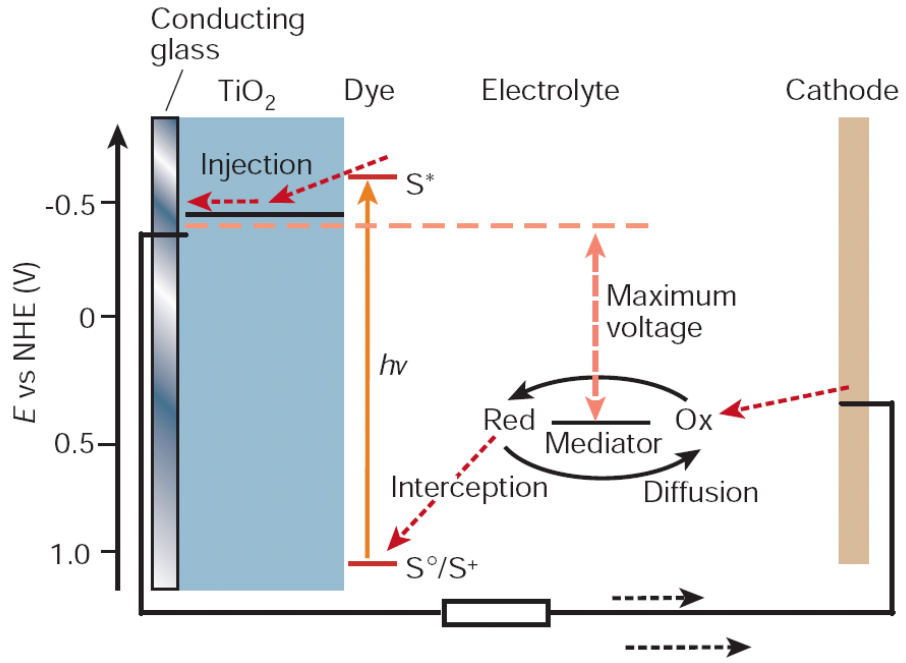

Figure 2.1: Schematic operation principle of dye sensitized solar cells[5].

to an unoccupied dye orbital, which energetically matches the conduction band of the semiconductor. After excitation the electron is injected into the conduction band of the semiconductor and diffuses from there through the semiconductor layers to the electrode. In the next step the electron passes an external load connecting the two electrodes, finally reaching the counter electrode. At the counter electrode the electron is transferred to the redox mediator, where it either reduces the oxidized form of the redox couple or fills a hole in the solid state electrolyte. Simultaneously to this process, the dye, oxidized by the electron injection, is regenerated by the electrolyte, respectively the hole conducting material.
Comparing the light to electron conversion in DSSC and conventional solar cells, a substantial difference has to be highlighted. Contrary to silicon based solar cells in which an electron from the valence band of the semiconductor is excited into the conduction band, leaving a hole in the conduction band, the adsorbed dye is excited and injects the excited electron in the conduction band of the semiconductor. While the electron-hole recombination process in silicon cells can limit the efficiency, this type of recombination is not possible in DSSC's, since holes are created in the dye and not in the semiconductor. The electron backtransfer from the semiconductor to the dye is in DSSC is negligible, due to the different timescales for the recombination and the

rereduction of the dye by the electrolyte. The recombination process in traditional devices limits the minimum light intensity needed to produce current, while DSSC's can work even under low light conditions.
Another important difference due to these different mechanisms is the process of charge separation and electron transport in the semiconductor. SC solar cells require a pn-junction, which induces an electric field gradient, and thus leads charge separation after light absorption. In DSSC charge separation is achieved by the electron injection from a dye LUMO into the conduction band of the semiconductor, and thus no electric field is needed. Therefore the requirements on the purity and composition of the semiconductor are much lower in DSSC. In addition nanocrystals are much easier to produce than single crystal semiconductor, ideally needed in traditional devices. For this reasons the production cost of DSSC should be much lower and the energy invested in the production should be regained much faster than for traditional solar cells.

2.2 Sensitizing Dyes

The quality of a dye serving as sensitizer depends on many different properties. Ideally the dye absorbs all photons below a wavelength of 900 nm and injects the excited electron with a quantum efficiency of one. Furthermore the regeneration by a given electrolyte should be fast enough to avoid the backtransfer. For practical use, there are additional properties, like a stable absorption on the semiconductor surface, a high long term stability (surviving a high number of redox turnovers) and for economical reasons cheap in production. A stable adsorption of the dyes on the semiconductor is usually obtained by introducing either a carboxylate or a phosphonate groups to the dyes, which are able to form a covalent bond to the surface. Except of the low cost argument, fulfilling at least partially each of the properties above is substantial for a suitable sensitizing dye.
In general sensitizers are divided into inorganic and organic dyes. The majority of inorganic dyes are ruthenium polypyridyl complexes. Next to the Ru-dyes, osmium polypyridyl dyes, metal porphyrines and phtalocyanines have been studied as well. Up to now, the ruthenium polypyridyl complexes (see Fig. 2.2) match the requirements listed above the best. Ruthenium dyes have a high molecular extinction coefficient, absorb in a wide range of visible light and are able to efficiently inject electrons into the semiconductor. In addition, these dyes have a high thermal and chemical stability in their neat form. Most of the common Ruthenium dyes are derivatives of the cis-bis(4,4'-dicarboxy-2,2'-bipyridine) dithiocyanato ruthenium(II) dye,

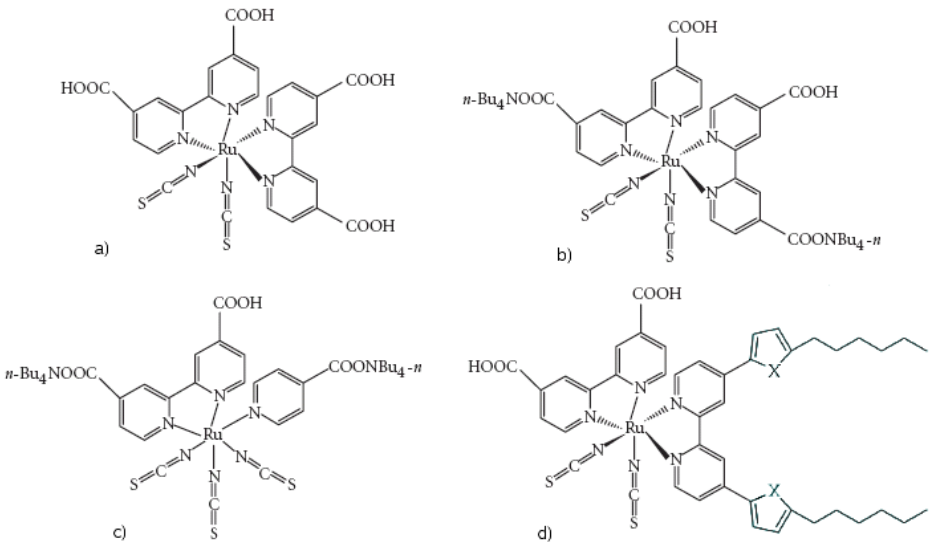

Figure 2.2: Structure of several high performant Ru-dyes. a) N3 dye, b) N719 dye c) black dye, c) C101 (X=O) and C102 (X=S)

better known as N3 dye. Mainly, modifications in the sidechains of the bipyridyl ligands are used to tune their performance. Other well performing dyes are the black dye (trithiocyanato (4,4,4-tricarboxy-2,2:6,2-terpyridine) ruthenium(II))[6] and its derivatives. Their properties are similar to the N3-like dyes, but their absorption spectrum is extended in the red region of the visible light. Contrary to the expectations, the wider absorption range of the terpyridine dyes in their current form does not lead to a significant improvement in the efficiency of the cells. The best dyes of the two types reach a similar performance of 11-12%.

The span of structural differences between organic sensitizers is much larger. The general structure of these dyes can be described as bridged conjugated $\pi-$donor $\pi-$acceptor groups. Examples therefore can be seen in the sensitizer presented in Fig 2.3. During the last years, organic dyes continuously improved, and a maximum efficiency of 9% was obtained with the D149 dye[7]. Organic sensitizers are more interesting for commercial cells, since they are, contrary to the Ruthenium dyes, cheap and easy to produce.

A common problem of all the dyes is their stability when being used in a DSSC device. Even being stable for many excitation-deexciation cycles in solution, the light induced oxidation of the dye at the semiconductor surface leads to instabilities. Therefore a fast regeneration of dyes with a more stable photooxidized product are required. In addition, all dyes are sensitive

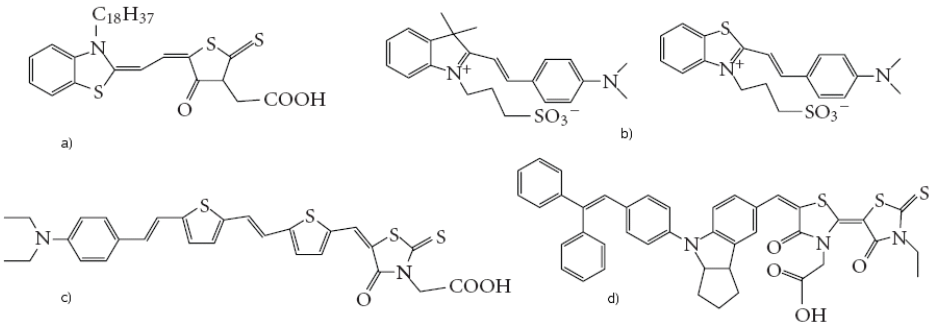

Figure 2.3: Common organic dyes used as sensitizers. a) merocyanine dye, b) hemicyanine dye, c) D-SS and d) Indoline dye D149

to molecular oxygen in DSSC. The exclusion of oxygen in DSSC is still a problem, since it requires an effective sealing with a high long term stability.

2.3 Electrolyte

The importance of the electrolyte in DSSC is manifold. The maximum voltage obtained by DSSC is determined by the energy difference between the conduction band of the semiconductor and the redox potential of the electrolyte (see Fig 2.1). Also the energy conversion efficiency and the stability of the cells depend on the electrolyte. Electrolytes used in DSSC are commonly divided into three classes: organic solvent electrolytes, ionic liquid electrolytes and solid electrolytes.
The redox couple leading to the highest efficiencies in organic and ionic liquid electrolytes is I^-/I_3^-. Nevertheless, only low voltages (≈ 0.7V) can be obtained with this redox couple. Literature reports several attempts to replace this couple either by different pseudohalid couples like $SCN^-/(SCN_2)$, $SeCN^-/(SeCN_2)$, to avoid the corrosive action of the I^-/I_3^- couple[8, 9]. Instead of using pseudohalids, other studies proposed inorganic complexes like substituted bipyridyl cobalt(II/III)[10] couples as a possible replacement. However, using these less aggressive substances significantly decreased the cell performance, thus I^-/I_3^- still represents the redox couple most widely used in organic and ionic liquid DSSC.
Organic solvents and ionic liquids have substantial differences influencing the efficiency and stability of the DSSC devices. Since charge transport occurs via the electrolyte, it has to provide a high ionic mobility for the ions used in the redox couple. Organic solvents have a low viscosity, which allows for a fast diffusion and therefore for a fast charge transport through the elec-

trolyte and a fast regeneration of the sensitizer. In addition a high particle exchange between bulk electrolyte and the pores of the nanocrystalline semiconductor film is required for an efficient regeneration of the dye molecules in this area. The organic solvents fulfilling these criteria the best are highly volatile short-chained nitrils like acetonitrile, propionitrile or valeronitrile. Except for these desirable properties, it has to be mentioned that all of these solvents are harmful to the environment. Unfortunately, a robust sealing of these cells, necessary to avoid contact with the environment, is hard to realize with organic solvents, since the high vapor pressure and thermal expansion at typical working temperatures put, next to the aggressive redox couple, additional stress on the sealing.

An alternative to organic solvents are ionic liquid electrolytes[11, 12, 13]. Even resulting in a lower energy conversion efficiency, they provide superior properties for the long term stability of DSSC devices. The advantages of ionic liquids are a negligible vapor pressure and a good chemical and thermal stability. Furthermore a high ionic conductivity and a high solubility of organic and inorganic materials make them a promising alternative.

Up to now, the most promising cations in ionic liquids are asymmetric N,N alkyl imidazolium ions. Typically the alkyl residues are short-chained linear residues (methyl, ethyl, propyl, butyl). The counter ion can either be I^-, $N(CN)_2^-$, $B(CN)_4^-$, BF_4^-, PF_6^- or SCN^-. By using different combinations of these ions, the properties of the ionic liquid can be easily tuned. The common problem of these liquids originates from their high viscosity, which slows down the I_3^- transport to the counter electrode and the ion exchange in the nanocrystalline semiconductor.

Solid state electrolytes are a relatively new class of electrolytes. Either a p-type semiconductor or a hole transporting material is used as electrolyte in this case. Up to now, there are several factors limiting the efficiency of these cells, like the slow charge transport in hole conductors, the infiltration of the nanocrystalline anatase network by the solid state electrolyte and stability issues. The maximum efficiency obtained with solid state DSSCs up to now is about 5%. Despite the weak performance, solid state DSSC are of high practical interest, since they provide desirable properties for the use in commercial applications and a high stability.

2.4 State of the Art and Future of DSSC Research

State of the art DSSC under laboratory conditions are able to compete with traditional cells, in terms of efficiency and stability. For DSSC an efficiency of more than 11% can be achieved using a single sensitizer. The use of cosensitizers in addition allows for efficiencies up to 15%. Under simulated conditions these cells lose less than 5% efficiency after 1000h at 80° Celsius in the dark, respectively 100h at 60° Celsius at 1 airmass (AM). Bigger modules, as used in practice, can be produced with 7% efficiency, which matches the lower end of conventional modules. In outdoor tests a loss of 15% in efficiency was detected after 4 years. The amount of electricity produced during this time is sufficient to regain the energy invested during production.
In order to systematically improve the efficiency, stability and applicability of DSSC, several key questions have to be answered. Most important for stability is the understanding of cell degradation and the underlying mechanisms. Several mechanisms have been proposed so far, but none of them consistently explains the degradation process. Therefore a detailed description of the physical and chemical processes induced by the interactions of the different substances in DSSC is necessary.
The efficiency of DSSC is highly sensitive to small changes in the cell assembly. For an improvement of the active side in DSSC the development of new sensitizing dyes is a key challenge. Since the dye represents the part of the active interface interacting the most with all other components, this development can not be separated from a deeper understanding of the physical and chemical effects induced by its environment.
The interactions responsible for stability and efficiency can be very specific and an atomistic model of the active regions in DSSC is important to guide the development of new devices. Next to experimental methods, computational chemistry can provide detailed insight. With increasing computer power and improved algorithms, theoretical chemistry can nowadays described processes on a scale of several nanometers, which is the expected range of the main interactions influencing the important processes at the active interface. In this work these techniques are used to provide an atomistic picture of the interfaces occurring in a high performant DSSC consisting out of an anatase semiconductor, the N3 dye and the acetonitrile based I^-/I_3^- electrolyte.

Chapter 3

Theory and Implementation

This chapter will give an overview of several methods commonly used in this work and new algorithms and methods implemented as part of this work. In case of the implemented methods, some implementation details and the performance will be discussed as well. The chapter starts with a brief introduction of density functional theory (DFT) and Born-Oppenheimer molecular dynamics (BOMD) as the basic ground state methods used in this work. In this context, the development of a multiple time step integrator for ab initio molecular dynamics using hybrid functionals will be discussed as well. Afterwards, time-dependent DFT (TDDFT) and Ehrenfest dynamics will be derived, and details of the implementation in CP2K will be presented. Furthermore, some remarks on mode selective vibrational analysis will be made.

3.1 Ground State Molecular Dynamics

3.1.1 Density Functional Theory

Nowadays, DFT is one of the most important methods in quantum chemistry to solve the static Schrödinger equation in the Born-Oppenheimer approximation,

$$\hat{H}\Psi = E\Psi, \tag{3.1}$$

with $\hat{H}$ being the Hamiltonian used for the quantum mechanical description of a given system. DFT provides the best balance between accuracy and computational cost for most systems of interest.
The idea of using functionals of the density to describe the properties of a quantum mechanical many body system was introduced by Thomas and Fermi in 1927. The theoretical proof of this concept was given in 1964 with the two Hohenberg-Kohn theorems.

- **Theorem 1:** The external potential v_{ext} is determined, within a trivial additive constant, by the electron density $\rho(\mathbf{r})$.

In other words the theorem states that the electron density can be used as basic variable to solve the Schrödinger equation, since it determines the ground state wavefunction and in this way all other electronic properties of an N-electron system. The proof of this theorem is trivial and involves only the minimum energy principle from wavefunction theory.

- **Theorem 2:** For a trial density $\tilde{\rho}(\mathbf{r})$ such that $\tilde{\rho}(\mathbf{r}) \geq 0$ and $\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = N$, it holds

$$E_0 \leq E_v[\tilde{\rho}] \tag{3.2}$$

  where E_0 is the exact ground state density and E_v is the energy functional defined as

$$E_v[\rho] = \int \rho(\mathbf{r}) v_{ext}(\mathbf{r}) d\mathbf{r} + F_{HK}[\rho]. \tag{3.3}$$

F_{HK} is a universal functional of the density that does not depend on v_{ext}, and collects the kinetic energy of the electrons and the electron electron interaction, which is composed out of the classical electron electron repulsion energy and a nonclassical term, namely the exchange-correlation energy.

$$F_{HK}[\rho]. = T_0[\rho] + J_{ee}[\rho] + E_{xc}[\rho]. \tag{3.4}$$

This theorem states that there is a variational principle in density functional theory analogue to the one known from wave function theory. The density for which the energy functional reaches the global minimum has to be the exact ground state density.
Since the exact form of F_{HK} is unknown, the accuracy of the theory strongly depends on the quality of the approximations to F_{HK}. The largest error in pure density functional theories, like the ones based on the Thomas-Fermi model, originates from approximations to the kinetic energy functional. Better approximations to the kinetic energy are obtained using the ansatz by Kohn and Sham developed in 1965. The Kohn-Sham ansatz [14] uses a different partitioning of Eq. 3.3 given by

$$E_{HK}[\rho] = T_s[\rho] + E_H[\rho] + E_{xc}[\rho] + \int \rho(\mathbf{r}) v_{ext}(\mathbf{r}) d\mathbf{r} \tag{3.5}$$

In this notation E_H is the electrostatic self energy of a classical charge distribution and T_s is the kinetic energy of a reference system of noninteracting electrons. Replacing the kinetic energy functional by the kinetic energy term known from wavefunction theory, T_s can be rewritten as

$$T_s[\rho] = \sum_i^N \left\langle \Psi_i \left| -\frac{1}{2}\nabla^2 \right| \Psi_i \right\rangle, \tag{3.6}$$

with Ψ_i being the one-electron wavefunction. The exchange correlation energy E_{xc} is then defined as $F_{HK} - T_s - E_H$, which is the collection of all contributions for which the exact functionals are unknown. Since Eq. (3.5) is only a repartitioning of Eq. (3.3), it is still valid to apply the variational principle from the 2nd Hohenberg-Kohn theorem, resulting in

$$\frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} + v_{KS}(\mathbf{r}) = \mu, \tag{3.7}$$

where μ is the Lagrangian multiplier associated to the constraint

$$\int \rho(\mathbf{r}) d\mathbf{r} = N, \tag{3.8}$$

which ensures the number of electrons N being conserved, and

$$v_{KS}(\mathbf{r}) = v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) + v_{ext}(\mathbf{r}). \tag{3.9}$$

The density of an interacting electron system in the external potential v_{ext} satisfying Eq.3.7 is the same as obtained for a noninteracting system in the external potential v_{KS}. For a given v_{ext} the ground state density can be obtained by solving the N one particle equations

$$\left[-\frac{1}{2}\nabla^2 + v_{KS}(\mathbf{r}) \right] \Psi_i = \epsilon_i \Psi_i \tag{3.10}$$

for the noninteracting system. The density is then computed as

$$\rho(\mathbf{r}) = \sum_i^N \sum_s |\Psi_i(\mathbf{r}, \tau_s)|^2, \tag{3.11}$$

with τ_s being the spin variable. The equations (3.10) are the Kohn-Sham equations. Each of the one particle equations in Eq. (3.10) is coupled to any other via the dependence of v_{KS} on $\rho(\mathbf{r})$, therefore Kohn-Sham equations have to be solved self-consistently. The Kohn-Sham ansatz allows the exact computation of the kinetic energy contribution of the noninteracting electrons T_s to the kinetic energy T_0. Since T_s is the major part of T_0 only a small fraction of T_0 has to be approximated within v_{xc}, which makes this approach superior to the ones directly approximating T_0.
Until now, the Kohn-Sham Ansatz is still an exact solution and even using wavefunctions to obtain an approximation for the kinetic energy, it remains a density functional theory. For applications, two approximations to the exact solution have to be made. The first is needed for the nonclassical electron

electron interactions and the leftover of the kinetic energy which are combined in the exchange correlation functional E_{xc}. The second approximation originates from the representation of the Kohn-Sham orbitals in a finite basis, in order to numerically solve the Kohn-Sham equations. Since the energy in the Kohn-Sham Ansatz is obtained variationally, a finite basis set always leads to an upper bound of the exact solution, but contrary to the error introduced in the xc-functional, the basis set error can be systematically reduced. An overview of several approximations to the exchange correlation functional and different classes of basis function and their properties will be given in the following sections.

3.1.2 Approximations to the Exchange and Correlation Functionals

The simplest approximation for $E_{xc}[\rho]$ is the local density approximation (LDA). In the LDA the assumption is made that the exchange-correlation energy of a inhomogeneous system is reasonably well described by the properties of the homogeneous electron gas of the same density. The energy is then given by

$$E_{xc}^{LDA}[\rho] = \int \rho(\mathbf{r})\varepsilon_{xc}(\rho)d\mathbf{r} \tag{3.12}$$

with the function

$$\varepsilon_{xc}(\rho) = \varepsilon_c(\rho) + \varepsilon_x(\rho) \tag{3.13}$$

consisting of the individual contributions of exchange functional and the correlation functional of the homogeneous electron gas. For this case an analytical solution for the exchange functional is known and can be directly inserted. The analytical solutions for the correlation functional are only known in the limits of high and low electronic density. Therefore the analytical form for the situation in atoms and molecules has to be interpolated.
The LDA is only justified for systems with slowly varying densities. As this is not the case for molecular systems, LDA is only a poor approximation in most cases. A more advanced approximation is the generalized gradient approximation (GGA): Instead of only using the electron distribution, the gradient of the electron density and thus its local variations are taken into account. For inhomogeneous distributions, this leads to a better description of the system. The energy is then defined as

$$E_{xc}^{GGA}[\rho] = \int \rho(\mathbf{r})\varepsilon_{xc}(\rho, \nabla\rho)d\mathbf{r}. \tag{3.14}$$

The exact form of $\varepsilon_{xc}(\rho, \nabla\rho)$ is unknown, and there are various functional forms including the gradient approximation for the exchange-correlation functional. Even taking the variations of the density into account, GGA still present a semi-local approximation as the density is always evaluated in a single point of space. Next to GGA's, meta-GGA functionals exist taking even higher derivatives into account. Meta GGA's lead to slightly better results, but the dominating error at this stage originates from neglecting nonlocal effects.
Therefore another class of functionals has been developed, using a fraction of exact exchange from Hartree-Fock theory in the exchange functional. These functionals are not longer pure density functionals, as they have a neat wave-function theory component. The energy functional for hybrid functionals is generally written as

$$E_{xc}^{hybrid}[\rho] = E_{xc}^{DFT} + a(E_x^{exact} - E_x^{DFT}) \tag{3.15}$$

with a specifying the amount of exact exchange mixed in. As the Hartree-Fock exchange term has to be computed for hybrid functionals, the computational cost significantly increases compared to GGA's.

3.1.3 Basis Sets

As mentioned above the quality of the basis functions used to expand the Kohn-Sham orbitals is important for the accuracy of the method. In general Ψ_i is expanded as a linear combination of basis functions.

$$\Psi_i(\mathbf{r}) = \sum_\nu c_{i\nu} f_\nu(\mathbf{r}, \{\mathbf{R}\}), \tag{3.16}$$

with the molecular orbital (MO) coefficients $c_{i\nu}$ to be optimized during the self consistent variational procedure. For the $f_\nu(\mathbf{r})$ arbitrary functions can be used which are able to represent any other function in the limit of a complete basis (basis set limit).
A physically motivated choice for $f_\nu(\mathbf{r})$ for the solution of the Kohn-Sham equations are functions similar to hydrogen atomic orbitals. Therefore Slater functions

$$f_\nu^S(\mathbf{r}) = N_\nu^S x^{m_x} y^{m_y} z^{m_z} e^{-\zeta_\nu |\mathbf{r}-\mathbf{R_A}|}, \tag{3.17}$$

where N_ν^S and ζ_ν are constants predefined for every basis, are a suitable, as their functional form for m_x=m_y=m_z=0 matches the form of the hydrogen 1s orbital. m_x, m_y and m_z define the angular momentum quantum numbers of the orbitals, and $\mathbf{r}$, resp. $\mathbf{R}$ are the electron and nuclei positions. These basis functions are known to be a very good basis in terms of convergence with

respect to the basis set limit. In practice, Slater functions are only rarely used, because of their numerically unfavorable properties. Thus, instead of Slater functions, Gaussian functions are widely used. Their general from is given by

$$f_\nu^G(\mathbf{r}) = N_\nu^G x^{m_x} y^{m_y} z^{m_z} e^{(-\zeta_\nu |\mathbf{r}-\mathbf{R_A}|^2)}. \tag{3.18}$$

Gaussians have similar properties as Slater functions, but avoid the numerical problems. The main difference between Slater functions and Gaussians is their long range decay and their shape around $\mathbf{r} - \mathbf{R_A} = \mathbf{0}$. Slater functions have a cusp at this point, which is the correct solution for atomic and molecular orbitals. Gaussian functions are smooth in this region. Therefore, either a larger number of basis functions, or a basis containing linear combination of Gaussians (contracted Gaussians) has to be used in order to resemble the correct shape in this region. By using contracted Gaussians as basis functions instead of a larger number single Gaussians, the number of MO coefficients to be optimized in the variational procedure is reduced and thus the computational cost for solving the KS equations.
Another widespread choice for basis functions are plane waves. These functions originate from solid-state theory. Plane waves are mainly used in systems with periodic boundary conditions, since they are periodic by construction

$$f_G^{PW}(\mathbf{r}) = N^{PW} e^{i\mathbf{Gr}}. \tag{3.19}$$

The normalization factor N^{PW} for these functions is given as the square root of the inverse volume of the simulation cell. Another desirable feature of plane waves is, except of the periodicity, that they form an orthonormal set of basis functions.
Comparing plane waves with Gaussians, a difference has to be highlighted, which becomes important in the later discussion of forces in molecular dynamics. Using Gaussian basis sets, the basis functions implicitly depend on the positions of the nuclei. Plane waves on the other hand are originless functions, which means that the derivative of a plane wave with respect to nuclear positions vanishes.

3.1.4 Born-Oppenheimer Dynamics

Assuming that the ionic motion obeys classical dynamics, the Lagrange function of a many body system with N classical particles in the Born Oppenheimer approximation reads as:

$$L(\mathbf{R}, \dot{\mathbf{R}}) = \sum_I^N \frac{1}{2} M_I \dot{\mathbf{R}}_\mathbf{I} + U(\mathbf{R}), \tag{3.20}$$

with $U(\mathbf{R})$ containing the nuclei-nuclei interaction energy and the interaction energy of the nuclei with the electronic subsystem. In BO-MD the potential energy surface spanned by the electrons is given by the solution of the static Schrödinger equation in the Born-Oppenheimer approximation.

$$U(\mathbf{R}) = \sum_{A<B} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} + \min_{\{\phi_i\}} \left[E^{KS}(\{\phi_i\}, \{\mathbf{R}\})\right] \tag{3.21}$$

As the ionic motion enters the electronic equations only parametrically, the effect of the ionic motion on the electronic motion is neglected and the potential only depends on their positions. Furthermore, since these equations are solved iteratively at every point, the electrons remain in the ground state at every time, therefore the potential will always be the ground state potential at a given set of atomic coordinates. From the Lagrange function, the equations of motion (EOM) are then derived via the associated Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L(\mathbf{R}, \dot{\mathbf{R}})}{\partial \dot{\mathbf{R}}} = \frac{\partial L(\mathbf{R}, \dot{\mathbf{R}})}{\partial \mathbf{R}}. \tag{3.22}$$

Thus the EOM become

$$M\ddot{\mathbf{R}} = -\frac{\partial U(\mathbf{R})}{\partial \mathbf{R}}. \tag{3.23}$$

When evaluating the forces, special care has to be taken of the electronic energy. In the Kohn-Sham approach the electronic term has to be replaced by the extended energy function of the converged wavefunctions ϕ_i,

$$\min_{\{\phi_i\}} \left[E^{KS}[\{\phi_i\}, \{\mathbf{R}\}]\right] = \varepsilon^{KS} = E^{KS} + \sum_{ij} \Lambda_{ij} \left[\langle\phi_i|\phi_j\rangle - \delta_{ij}\right], \tag{3.24}$$

including the orthonormality constraint with the Lagrange multipliers Λ_{ij}. The electronic forces are obtained by taking the derivative of this expression with respect to $\mathbf{R}$.

$$\frac{d\varepsilon^{KS}}{d\mathbf{R}} = \frac{\partial E^{KS}}{\partial \mathbf{R}} + \sum_{ij} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}} \left[\langle\phi_i|\phi_j\rangle - \delta_{ij}\right] + \sum_i \left[\frac{\partial E^{KS}}{\partial \langle\phi_i|} + \sum_j \Lambda_{ij}|\phi_j\rangle\right] \frac{\partial \langle\phi_i|}{\partial \mathbf{R}} \tag{3.25}$$

At the variational minimum, the term in brackets vanishes, and equation 3.25 reduces to

$$F^{KS}(\mathbf{R}) = \frac{\partial E^{KS}}{\partial \mathbf{R}} + \sum_{ij} \Lambda \frac{\partial}{\partial \mathbf{R}} \langle\phi_i|\phi_j\rangle. \tag{3.26}$$

The last term of this equation is a derivative of the wavefunctions with respect to the ionic positions. Therefore, it remains only in case basis functions are used which have an explicit dependence on the ionic position, like Gaussian or Slater functions.

3.2 Multiple Time Step Molecular Dynamics

Common integration schemes for Born-Oppenheimer MD, like the velocity Verlet algorithm, allow only for timesteps that are a fraction of the theoretical limit. The latter is given by half the frequency of the fastest motion in the system[15]. In case hybrid functionals are employed for the electronic part, the four center integrals, which are the most expensive part of these calculations have to be recalculated at each step. Therefore, an algorithm allowing for larger timesteps is especially useful for molecular dynamics based on hybrid functionals. As part of this work a multiple time step algorithm combining local functionals and hybrids has been developed.

3.2.1 The Multiple Time Step Integrator

The DFT multiple time step (DFT-MTS) integrator presented here is based on the reversible reference system propagator algorithm [16] (r-RESPA), known from force field based classical dynamics. Using a carefully constructed integration scheme, the time evolution remains reversible, and the molecular dynamics simulation remains accurate and energy conserving. In DFT-MTS, the difference in computational cost between a hybrid and a local functional is exploited, by performing a hybrid calculation only after several pure DFT calculations. r-RESPA is derived from the Liouville operator representation of Hamiltonian mechanics

$$iL = \sum_{j=1}^{f} \left[\frac{\partial H}{\partial p_j} \frac{\partial}{\partial x_j} + \frac{\partial H}{\partial x_j} \frac{\partial}{\partial p_j} \right], \tag{3.27}$$

where L denotes the Liouville operator for the system containing f degrees of freedom. This operator is then used to create the classical propagator $U(t)$ for the system:

$$U(t) = e^{iLt}. \tag{3.28}$$

Decomposing the Liouville operator into two parts

$$iL = iL_1 + iL_2, \tag{3.29}$$

and applying a 2nd-order Trotter-decomposition to the corresponding propagator yields

$$e^{i(L_1+L_2)\Delta t} = \left[e^{i(L_1+L_2)\Delta t/n}\right]^n = \left[e^{iL_1(\delta t/2)} e^{iL_2 \delta t} e^{iL_1(\delta t/2)}\right]^n + O(\delta t^3), \tag{3.30}$$

with $\delta t = \Delta t/n$. For this propagator several integrator schemes can be derived[17]. The extension for multiple time step (MTS) MD is obtained

from a decomposition of the force in the Liouville operator into two or more separate forces

$$iL = \sum_{j=1}^{f} \left[\dot{x}_j \frac{\partial}{\partial x_j} + F_j^1 \frac{\partial}{\partial p_j} + F_j^2 \frac{\partial}{\partial p_j} \right]. \tag{3.31}$$

For that specific case, the propagator reads

$$e^{iL\Delta t} = e^{(\Delta t/2)F^2 \frac{\partial}{\partial p}} \left[e^{(\delta t/2)F^1 \frac{\partial}{\partial p}} e^{\delta t \dot{x}_j \frac{\partial}{\partial x_j}} e^{(\delta t/2)F^1 \frac{\partial}{\partial p}} \right]^n e^{(\Delta t/2)F^2 \frac{\partial}{\partial p}}. \tag{3.32}$$

This allows to treat F^1 and F^2 with different time steps, while the whole propagator still remains time reversible. The procedure for F^1 and F^2 will be referred to as the inner and the outer loop, respectively. In the DFT-MTS approach, the forces are split in the following way

$$\begin{aligned} F^1 &= F^{\text{local}} \\ F^2 &= F^{\text{hybrid}} - F^{\text{local}} \end{aligned}$$

where F^{hybrid} are the forces as obtained from a hybrid calculation, and F^{local} the forces as obtained from a local functional. It is obvious that the corresponding Liouville operator equals a purely hybrid one. The advantage of this splitting is that the magnitude of F^2 is usually much smaller than of F^1 or F^{hybrid}. To appreciate that, it has to be considered how closely geometries and frequencies obtained by a hybrid functional normally match the ones obtained by a local functional, in particular for stiff degrees of freedom. The difference of the corresponding Hessians is therefore small and low-frequent. However, high-frequency parts are not removed analytically, thus the theoretical upper limit for the time step of the outer loop remains half the period of the fastest vibration[15]. The gain originates from an increased accuracy and stability for larger time steps in the outer loop integration. Even using an outer loop time step close to the theoretical limit, a stable and accurate MD is obtained. In particular, contrary to the single time step case, there is no shift to higher frequencies as the (outer loop) time step is increased.

3.2.2 Validation and Benchmarks

In order to validate the correctness of the MTS-scheme, several MD simulations on a water cluster containing 6 water molecules have been performed. The inner loop used the PBE functional and a 0.5fs time step, while in the outer loop the PBE0 hybrid functional with various time steps was employed. These are referred to as MTS-n, where n indicates the the length

of the outer loop time step (n × 0.5fs). Shown in Fig. 3.1 is the constant of motion (total energy) as obtained during a few ps of MTS MD. Stable molecular dynamics is obtained up to n=7 or n=8. The onset of instability

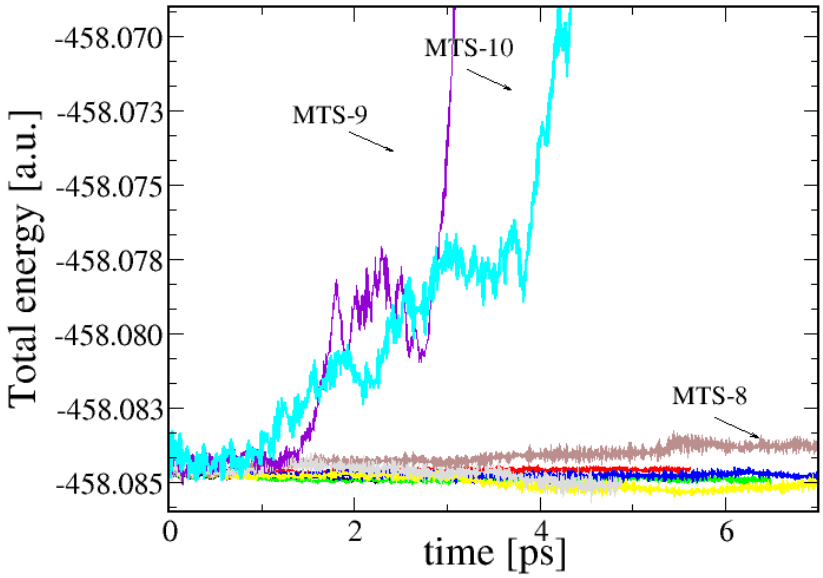

Figure 3.1: Constant of motion (total energy) for a cluster containing 6 water molecules during MTS MD with various time steps (see text for details). The curves for MTS-2 to MTS-6 are all stable and lie almost on top of each other, and are therefore not separately named. With MTS-8 the beginning of an instability can be seen. MTS-9 and MTS-10 lead to unstable simulations for this system.

corresponds approximately to a time step that equals half the period of the OH stretch. Despite the fact that a time step of 2.5fs is unstable in single time step MD, MTS MD with an outer loop time step of 2.5fs ($n = 5$) is a perfectly reasonable choice. Next to energy conservation, the vibrational frequency of the OH stretching mode remains unaffected by the MTS scheme up to n=7. This indicates the quality of this algorithm, as a shift of 200 cm^{-1} is already seen at a timestep of 1.5fs in single time step (STS) MD. For further validation and benchmarking liquid water has been calculated once using the MTS-5 scheme resulting in a timestep of 2.5fs, and twice (HSE and PBE) with the normal STS scheme employing a timestep of 0.5fs. The simulation cell contained 64 water molecules, using a triple-ζ basis and two

polarization functions for hydrogen and oxygen. All three calculations start from the same point in phase space, and the time evolution of the potential energy is monitored. These results show clearly that the MTS-5 run follows closely the HSE trajectory, and not the PBE trajectory, despite the fact that the MTS-5 run only evaluates the HSE forces once for every five evaluations of the PBE force. Comparing the timings on 32 CPUs for this system, a 3.6 fold increase in simulation speed is obtained when applying the MTS-5 scheme (conventional HSE 1120s per fs of MD, HSE MTS-5 308 s per fs of MD). Using higher order MTS scheme like MTS-7 leads to similar results, but the speed up is not significant, as the wavefunction extrapolation gets less efficient. Hence more wavefunction optimization steps in the hybrid part are required, which compensate the gain of the larger timestep. From the discussion above it becomes obvious that the DFT-MTS scheme allows for faster hybrid functional MD, without loss in accuracy.

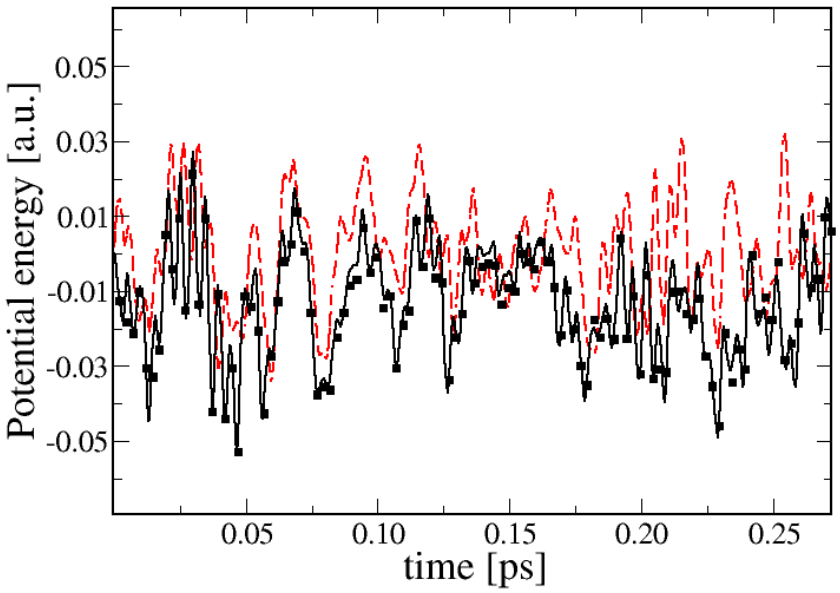

Figure 3.2: Shown are the potential energies for a bulk system of liquid water during MD as obtained with PBE (red dashed line), HSE (black solid line) and HSE with the MTS-5 scheme (squares). All simulations were started from the same point in phase-space. To simplify comparison, PBE and HSE energies have been aligned at t=0.

3.3 Quantum-Classical Molecular Dynamics

3.3.1 Time Dependent DFT

The previous section described DFT for ground state calculations. These methods are not able to describe excited state properties or electron dynamics. Therefore time-dependent DFT is needed. Instead of solving the static Schrödinger equation it aims to solve the time dependent Schrödinger equation (TDSE),

$$i\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \hat{H}(\mathbf{r},t)\Psi(\mathbf{r},t) \tag{3.33}$$

Contrary to the time independent case, there is no variational principle for the total energy. As a consequence the theorems of Hohenberg and Kohn do not hold for that case. The extension for time dependent systems is the Runge-Gross theorem[18], which states

- **Runge-Gross Theorem:** There is a one-to-one correspondence between the external potential $v_{ext}(\mathbf{r},t)$ and the electronic density $\rho(\mathbf{r},t)$, for systems evolving from a fixed many-body state.

This theorem has the same meaning as the first HK-theorem, but presents a generalization for time dependent systems. The proof of the Runge-Gross theorem is somewhat more complicated than for the HK-theorem. At first it has to be shown that external potentials differing by more than a time dependent function, produce different current densities. In the second step it has to be verified that different current densities imply different electron densities. Thus the external potential uniquely defines the electron density. As there is no variational principle, the total energy has to be replaced by the quantum mechanical action,

$$A[\Psi] = \int_{t_0}^{t_1} dt\langle\Psi(t)|i\frac{\partial}{\partial t} - \hat{H}(t)|\Psi(t)\rangle. \tag{3.34}$$

for which it is valid that the function $\Psi(t)$ which makes the action stationary will be the solution of the TDSE. In addition, at the solution point $\Psi(t)$ the action will always be zero.
The findings above allow to formulate a Kohn-Sham like ansatz for time dependent systems. The resulting time-dependent Kohn-Sham (TD-KS) equations read as

$$i\frac{\partial}{\partial t}\psi_i(\mathbf{r},t) = \left[-\frac{\nabla^2}{2} + v_{KS}(\mathbf{r},t)\right]\psi_i(\mathbf{r},t) \tag{3.35}$$

with ψ_i the Kohn-Sham orbitals and the time-dependent density defined as

$$\rho(\mathbf{r},t) = \sum_{i}^{\text{occ}} |\Psi(\mathbf{r},t)|^2. \tag{3.36}$$

Comparing this equation and its time independent form, there is a substantial mathematical difference between the two differential equations: The time independent one is a boundary value problem, while the time-dependent is an initial value problem. Hence, the time-dependent KS-equations can not be solved iteratively.

3.3.2 Propagators for TDDFT

A practical method for solving the TDKS equation is the so called real time propagation. For a given initial state at time t_0, the wavefunctions, respectively the density, of a system are propagated stepwise to a final state at t_f[19]. In order to derive a numerically feasible method for the propagation, the TDKS equation is reformulated in its integral form

$$\psi(\mathbf{r},t_f) = \hat{U}(t_f,0)\psi(\mathbf{r},t_0) \tag{3.37}$$

with the time evolution operator $\hat{U}$

$$\hat{U}(t,0) = \hat{T}\exp\left\{-i\int_0^t H(\tau,\mathbf{r})d\tau\right\} \tag{3.38}$$

and $\hat{T}\exp$ as the time ordered exponential. The time evolution operator (TEO) in the exact reformulation of the TDKS equation is a short notation for

$$\hat{U}(t,0) = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\int_0^t d\tau_1\int_0^t d\tau_2\ldots\int_0^t d\tau_n\hat{T}\{H(\tau_1)H(\tau_2)\ldots H(\tau_n)\}. \tag{3.39}$$

For obvious reasons the time evolution operator is not suitable for real applications. In general, the state at t_f is not directly derived from t_0. Instead the interval is split into smaller timesteps δt, for which the propagation is performed. Using small time steps reduces the time dependences of the Hamiltonian, therefore a single exponential approximation can be made to the time evolution operator.
There are two important properties of the time evolution operator which should be retained by the approximated propagators. The propagators have

to be unitary, which is equivalent to conserving the total charge during propagation. In addition to unitary, the propagators should be time reversible, which is important for the stability of the propagation. As an example of a time reversible and unitary propagator the "enforced time-reversible symmetry" propagator (ETRS) will be discussed. A more detailed discussion on this topic can be found in [19]. The derivation of the ETRS propagator starts with an exponential approximation to the TOE.

$$\hat{U}(t,0) = \exp\{-it\hat{H}(0)\}, \tag{3.40}$$

For a time reversible propagator it has to be valid that propagating $\psi(t)$ by $\delta t/2$ forwards has to give the same result as the backwards propagation of $\psi(t+\delta t)$ by $\delta t/2$. This condition reads as

$$\exp\left\{+i\frac{\delta t}{2}\hat{H}(t+\delta t)\right\}\psi(t+\delta t) = \exp\left\{-i\frac{\delta t}{2}\hat{H}(t)\right\}\psi(t). \tag{3.41}$$

Rearranging this equations leads to the ETRS propagator

$$U_{ETRS}(t+\delta t,t) = \exp\left\{-i\frac{\delta t}{2}\hat{H}(t+\delta t)\right\}\exp\left\{-i\frac{\delta t}{2}\hat{H}(t)\right\}. \tag{3.42}$$

In general, the ETRS propagator is computed self consistently. For the self consistent operator it can be shown, that it is not only time reversible but also unitary. In section 3.4 some details about the implementation of different propagators will be given.

3.3.3 Ehrenfest Dynamics

A rather general derivation for Ehrenfest (non-adiabatic quantum-classical molecular) dynamics can be obtained starting from the action of a system [20]. For a situation in which the the electrons are treated quantum mechanically while the nuclei are treated classically the total action can be written as the sum of the two environments

$$A_{\text{tot}} = A_c + A_q \tag{3.43}$$

where the classical action of the nuclei is defined as

$$A_c = \int_{t_0}^{t_1}\left[\sum_{A}^{N_i}\frac{M_A}{2}\dot{R}_A^2 - U(\mathbf{R},t)\right]dt \tag{3.44}$$

and the quantum mechanical action as in Eq. 3.34. The equations of motion are then derived by making the action stationary, i.e.

$$\frac{\delta A}{\delta \langle \Psi(\mathbf{r},t)|} = 0 \tag{3.45}$$

and

$$\frac{\delta A}{\delta \mathbf{R}(t)} = 0 \tag{3.46}$$

Evaluating these expressions in the framework of TDDFT, the equations of motion become

$$M\ddot{\mathbf{R}} = -\frac{\partial}{\partial \mathbf{R}} U(\mathbf{R},t) - \sum_j \left\langle \Psi^j \left| \frac{\partial}{\partial \mathbf{R}} V_{\text{int}}(\mathbf{r},\mathbf{R}) \right| \Psi^j \right\rangle \tag{3.47}$$

for the nuclear motion, and

$$i\frac{\partial}{\partial t}\Psi^j = \hat{H}(\mathbf{r},\mathbf{R},t)\Psi^j \tag{3.48}$$

for the electrons, which is exactly the time dependent Schrödinger equation solved by TDDFT. These equations are valid for the exact wavefunctions. In the Kohn-Sham approach, the wavefunctions are replaced by a linear combination of basis functions. For plane waves the equations remain the same, since they do not depend on the nuclear coordinates and therefore are independent of the change of the ionic positions during MD. The case of a Gaussian basis set requires a more detailed analysis, which will be presented in the next section.

3.3.4 Ehrenfest Dynamics Using Gaussian Basis Sets

Using Gaussian basis sets for the expansion of the wavefunctions an implicit dependence of the wavefunctions on the nuclear positions is introduced. Since in an MD approach the nuclear positions are function of time, the time derivative in the quantum mechanical action

$$A[\Psi] = \int_{t_0}^{t_1} dt \sum_j \langle \sum_\alpha c_\alpha^j(t)\phi(\mathbf{r}-\mathbf{R}_A)|i\frac{\partial}{\partial t} - \hat{H}(t)|\sum_\alpha c_\alpha^j(t)\phi(\mathbf{r}-\mathbf{R}_A)\rangle \tag{3.49}$$

has to be replaced by the total time derivative

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \sum_A \frac{\partial \mathbf{R}_A}{\partial t}\frac{\partial}{\partial \mathbf{R}_A}. \tag{3.50}$$

The EOM can be derived the same way as described in the previous section, except that due to the introduction of the basis the independent variables for making the action constant become the coefficient $c_\alpha(t)$ instead of the wavefunction. This leads to the EOM for the electronic subsystem

$$\dot{c}^j_\alpha = -\sum_{\beta\gamma} S^{-1}_{\alpha\beta}(iH_{\beta\gamma} + B_{\beta\gamma})c^j_\gamma \tag{3.51}$$

with the definitions

$$S_{\alpha\beta} = \langle\phi_\alpha|\phi_\beta\rangle \qquad \text{and} \qquad B_{\alpha\beta} = \left\langle \phi_\alpha|\frac{d}{dt}\phi_\beta \right\rangle. \tag{3.52}$$

Equation (3.51) has the same structure as the TDKS equations. For solving this equation the same propagation techniques can be used as for the case of fixed ionic positions (real time propagation). In Ehrenfest dynamics an additional contribution due to the derivative of the basis functions becomes part of the Hamiltonian used in the exponential operator. Instead of being purely imaginary, the matrix in the exponential of the propagator becomes fully complex.
Due to the dependence of the basis function on the nuclear positions, special care has to be taken of the functional derivative with respect to $\mathbf{R}(t)$ as well. Using Euler's equation, after some algebra, the EOM for the nuclei are obtained as

$$M_A\ddot{\mathbf{R}}_A = -\frac{\partial U(\mathbf{R},t)}{\partial \mathbf{R}_A} + \sum_j^{N_e}\sum_{\alpha\beta} c^{j*}_\alpha \left(\frac{\partial H_{\alpha\beta}}{\partial \mathbf{R}_A} + D^A_{\alpha\beta}\right) c^j_\beta \tag{3.53}$$

with $D^A_{\alpha\beta}$ collecting all terms due to the dependence of the wavefunction on $\mathbf{R}_A$,

$$D^A_{\alpha\beta} = \sum_{\gamma\delta}(B^{A+}_{\alpha\gamma}S^{-1}_{\gamma\delta}H_{\delta\beta} + H_{\alpha\gamma}S^{-1}_{\gamma\delta}B^{A+}_{\delta\beta}) + \tag{3.54}$$

$$i\left[C^{A+}_{\alpha\beta} - C^A_{\alpha\beta} + \sum_{\gamma\delta}(B^+_{\alpha\gamma}S^{-1}_{\gamma\delta}B^A_{\delta\beta} + B^{A+}_{\alpha\gamma}S^{-1}_{\gamma\delta}B_{\delta\beta})\right], \tag{3.55}$$

with

$$B^A_{\alpha\beta} = \left\langle \phi_\alpha|\frac{\partial}{\partial \mathbf{R}_A}\phi_\beta \right\rangle \qquad \text{and} \qquad C^A_{\alpha\beta} = \left\langle \frac{d}{dt}\phi_\alpha|\frac{\partial}{\partial \mathbf{R}_A}\phi_\beta \right\rangle. \tag{3.56}$$

All Terms collected in $D^A_{\alpha\beta}$ only appear when atom centered basis sets are used.

3.4 Implementation of TDDFT and Ehrenfest dynamics

3.4.1 General remarks

The implementation of Real Time Propagation and Ehrenfest MD was done in the framework of the CP2K program package. For the calculation of the Kohn-Sham matrix the existing modules were used. New routines had to be written for the calculation of the different derivative matrices B,B^A and C^A. Three different propagators have been implemented, the ETRS propagator as mentioned above, the exponential midpoint (EM) propagator

$$\hat{U}_{EM}(t+\delta t,t) = \exp\{i\delta t\hat{H}(t+\delta/2)\} \tag{3.57}$$

and the Crank-Nicholson propagator which can be seen as a first order Padé approximation of the EM propagator. In the real time propagation scheme (fixed ionic positions) the self consistent solution only involves the calculation of the new Kohn-Sham matrix for the propagated coefficients. For Ehrenfest MD, the iterative procedure is embedded into the integrator of the nuclear EOM. This is a result of the velocity dependence of the B matrix defined in 3.52 which has to be added to the Hamiltonian when Gaussian basis sets are used. Hence, each iteration step for Ehrenfest dynamics involves a real time propagation step and a complete evaluation of the forces. Due to this, more iterations are needed to reach self consistency for the propagator. The convergence criterium implemented in CP2K is defined as

$$||\Delta C^T S \Delta C||_{\max} \leq \epsilon \tag{3.58}$$

with ΔC as the difference of coefficient matrices in two successive steps.

3.4.2 Matrix Exponentials

In terms of computational cost, the most expensive part is the evaluation of the matrix exponential in the propagator. A list of methods how to compute the exponential of a matrix is given in [21]. Four different methods have been implemented in CP2K, which will be discussed in detail.

Method 1: Taylor expansion

The formal solution for calculating a matrix exponential, is the expansion of the exponential in an infinite Taylor series.

$$\exp(A) = \text{Taylor}\{A\} = \sum_{n=0}^{\infty} \frac{1}{n!} A^n. \tag{3.59}$$

In many applications this series is truncated after 4th or 6th order. Truncated Taylor expansions are only conditionally stable. If the 2-norm of the matrix is small enough the series converges. This can be achieved by using a small value of δt in the propagator. For such a case the series can be evaluated more efficiently, successively calculating the matrix vector products

$$e^A v = \text{Taylor}\{A, v\} = \sum_{n=0}^{\infty} \frac{1}{n!} A^n v \ . \tag{3.60}$$

This method is especially useful in a plane-wave basis, since the dimension of A is rather large in general.
For a code using Gaussian basis sets (where the dimension of A is usually smaller), a stable Taylor expansion can be achieved by a scaling and squaring procedure

$$e^A = \left[\exp\left(\frac{A}{2^m}\right)\right]^{2^m} = \left[\sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{A}{2^m}\right)^n\right]^{2^m} \ . \tag{3.61}$$

Using this procedure the matrix norm is decreased by a factor of 2^m for the cost of m additional matrix multiplies. This procedure allows for much larger timesteps, which compensate the additional cost of the squaring.

Method 2: Diagonalization

For a matrix, which can be diagonalized using a unitary transformation

$$U^T A U = A_D \qquad \text{with} \qquad U^T U = 1 \tag{3.62}$$

the Taylor series becomes

$$e^A = U \text{Taylor}\{A_D\} U^T = U e^{A_D} U^T. \tag{3.63}$$

Thus, the problem reduces to computing the exponential of the diagonal matrix and backtransforming this result. Beside being the most accurate way of calculating the matrix exponential, this method suffers from several disadvantages.

- The matrix has to be diagonalizable
- Not feasible for large system, because of the O(N^3)-scaling
- Bad parallel performance
- No library routines for general complex matrices.

For these reasons, the method is not suitable for real simulations, but can be used for optimizing the accuracy needed for the other methods discussed.

Method 3: Padé approximation

Another method using a polynomial expansion for a matrix exponential is the Padé approximation

$$\exp(A) = D_{pq}^{-1}(A) N_{pq}(A) \tag{3.64}$$

with

$$N_{pq}(A) = \sum_{j=0}^{p} \frac{(p+q-j)!p!}{(p+q)!j!(p-j)!} A^j \tag{3.65}$$

and

$$D_{pq}(A) = \sum_{j=0}^{q} \frac{(p+q-j)!q!}{(p+q)!j!(q-j)!} A^j. \tag{3.66}$$

Using this method $N_{pq}(A)$ and $D_{pq}(A)$ have to be evaluated first, before the exponential can be calculated. This makes the stepwise matrix vector product as applied in the Taylor expansion impossible. For this reason, it is only rarely used in other codes. The Padé approximation still allows a scaling and squaring procedure, as introduced above. Combined with the faster convergence and increased stability, using the Padé approximation is more efficient in this scheme than the Taylor expansion. Both polynomial expansions are valid for general complex matrices, thus they can be used for Ehrenfest dynamics and real time propagation.

Method 4: Arnoldi subspace iteration

The Arnoldi method is a generalization of the Krylov subspace projection, which can be used for general complex matrices. The Krylov space for a given matrix A and a vector v is defined as

$$K_m\{A, v\} = \text{span}\{v, Av, A^2v, ..., A^{m-1}v\} \tag{3.67}$$

Rewriting the Taylor expansion in terms of the action on the vector v

$$e^A v = c_0 v + c_1 A v + c_2 A^2 v + ... + c_{m-1} A^{m-1} v \tag{3.68}$$

it becomes clear that this is an element of the Krylov space $K_m\{v, A\}$. Furthermore, all other polynomial expansions up to order $m-1$ are elements of this space, and thus the optimal approximation as well.
There are several advantages, which will become obvious later on, in constructing the space in an orthonormal basis. For general matrices, this basis can be build using the Arnoldi procedure

$\beta = ||v||_2$
$v_1 = \frac{v}{\beta}$
`DO i=1,m`
  $p = A v_i$
  `DO j=1,i`
    $h_{ji} = v_j \ p$
    $p = p - h_{ji} v_j$
  `END DO`
  $h_{j+1,j} = ||p||_2$
  $v_{j+1} = \frac{p}{h_{j+1,j}}$
`END DO`

The resulting orthonormal basis vectors can be collected in a matrix $V_m = \{v_1, v_2, ..., v_m\}$. In addition, The projection of A on V is directly obtained by the element h_{ij} (elements of matrix H). The connection between the matrices is given by

$$A V_m = V_{m+1} \tilde{H} = V_m H_m + h_{m+1,m} v_{m+1} e_m^T, \tag{3.69}$$

with e_m being the m-th unit vector. Since the basis V is orthonormal by definition, it follows

$$V_m^T V_m = 1 \qquad \text{and} \qquad V_m^T v_{m+1} = 0 \tag{3.70}$$

Thus by multiplying equation 3.69 with V_m^T

$$V_m^T A V_m = H, \tag{3.71}$$

is obtained. It can be shown that the best fit to $e^A v$ in the orthonormal Krylov subspace is given by

$$e^A v \approx \beta V_m e_m^H e_1. \tag{3.72}$$

Since H is a square matrix with the dimension of the Krylov space (usually <20), the exponential can be computed efficiently using one of the methods

described above. The computational cost of this procedure is dominated by the construction of the Krylov space, more precisely by the calculation of the matrix vector products Av. For the propagation this procedure has to be applied for every MO vector, therefore the computational cost grows with $O(n_{ao}^2 n_{mo})$, with n_{ao} is the number of basis functions and n_{mo} the number of molecular orbitals.

Performance

The test on the parallel scalability and the computational cost was made on a system containing 15733 basis function and 2944 molecular orbitals. For systems of this size, the computation of the matrix exponential becomes the dominant contribution to the computational time needed. For these test calculations 6 propagation steps have been computed, each requiring 6 iterations.
In this benchmark different implementations of the Arnoldi method have been tested. The first one was a straightforward implementation of the Arnoldi method. In the second one, the code was optimized for parallel scalability, by rearranging matrices and tuning the SCALAPCK matrix block sizes. Furthermore, the number of matrix exponential calculations could be reduced by exploiting the fact, that in two successive MD steps using the ETRS scheme the Arnoldi subspace can be reused. For the Taylor and the Padé method no significant speedup could be achieved with such changes.
From Figure 3.3 it becomes obvious that the Arnoldi method provides a superior performance. Comparing the theoretical scaling, the Arnoldi method is expected to be about 5 times as fast as Padé or Taylor. The real speed up is roughly a factor of 3-4. This is due to the different number of matrix matrix multiplications needed in the different methods. In case of Taylor and Padé approximation, the number of matrix multiplications is optimized using an estimate of the 2-norm of the initial Kohn-Sham matrix. As the Arnoldi method is an iterative procedure, an adaptive algorithm can be applied, and therefore, the number of steps can vary. For matrices with a large 2-norm, the number of Arnoldi steps, and thus matrix multiplications exceeds the number of multiplications needed in the other schemes. In the test case, Padé required 7 multiplications and a single inversion, while 13 Arnoldi steps were needed to reach convergence. For systems in which the Kohn-Sham matrix have even larger 2-norms (e.g. when using larger time steps), the Pade approximation can be the faster and more stable choice than the Arnoldi method.

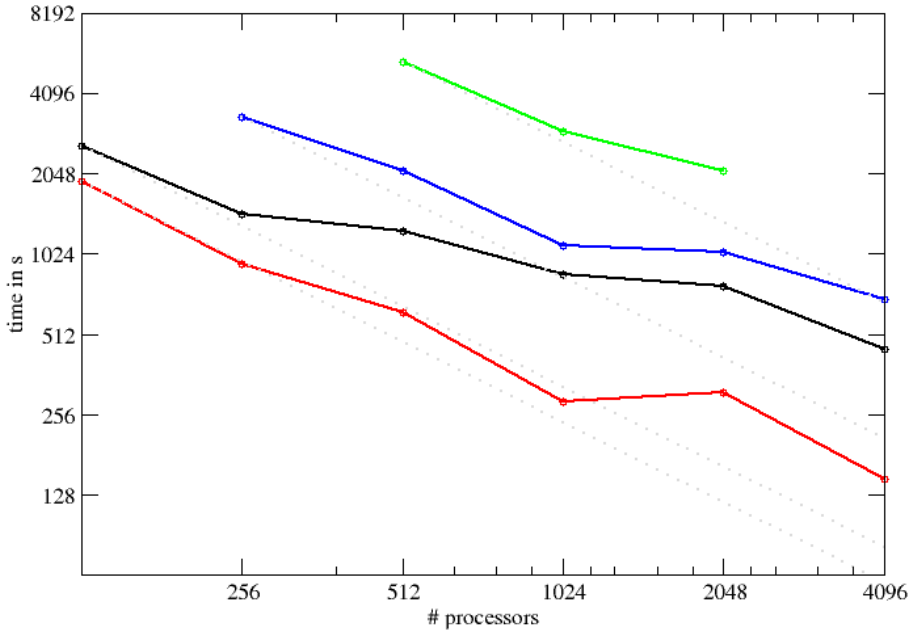

Figure 3.3: Timings and parallel scaling for real time propagation using different matrix exponentials. The lines are unoptimized Arnoldi(black), optimized Arnoldi(red), Padé approximation (blue), Taylor expansion (green). The dotted lines indicates the optimal scaling for each method.

3.4.3 Extrapolation

The propagation schemes discussed in section 3.3.2 require an iterative procedure. For these type of problems it is common to apply extrapolation schemes in order to speed up convergence. The extrapolation scheme implemented is the always stable predictor corrector (ASPC) [22] method.
As implemented in CP2K, several quantities can be used for extrapolation. In the real time propagation scheme an extrapolation on the MO coefficients is possible. For Ehrenfest dynamics, it turned out to be more useful to perform an extrapolation based on the product of the density and the overlap matrix.

$$C(t_n) \approx \sum_{i=1}^{n-1} d_i C(t_{n-i}) C(t_{n-i})^T S(t_{n-i}) C(t_{n-1}) \tag{3.73}$$

For methods 1-3 in section 3.4.2 the extrapolation can be applied directly on the exponential as well. Method 4 does not allow for extrapolating the exponential, instead the matrix to be exponentiated (Kohn-Sham matrix)

can be extrapolated itself. Hence, using the Arnoldi method, requires an additional evaluation of the matrix exponential in the first iteration step. The quality of the different extrapolation strongly depends on the system applied to and the settings of the simulation. Therefore, which method to use and the extrapolation order needs to be tested on the system of interest.

3.5 Vibrational Analysis

3.5.1 Normal Mode Analysis

Infrared vibrational spectroscopy is a widely used tool to characterize structural properties of a substance. In computational chemistry, a quantum mechanical description of these properties is often obtained using normal mode analysis (NMA). The theoretical description within NMA is obtained from the nuclear Schrödinger equation in the Born-Oppenheimer approximation.

$$\left[T_n + E_{e,n}(\mathbf{R})\right] \Phi_A \mathbf{R} = E_{\text{tot}} \Phi_A \mathbf{R} \tag{3.74}$$

In this equation $\Phi_A \mathbf{R}$ is the nuclear wavefunction, T_n the kinetic energy operator of the nuclei and $E_{e,n}(\mathbf{R})$ the energy of the electronic subsystem for a given set of nuclear position $\mathbf{R}$. A common ansatz for the nuclear wavefunction is a product of functions depending only on a single nuclear coordinate. Therefore, the individual coordinates have to be decoupled. A decoupling can be achieved by expanding the potential energy surface spanned by the electronic subsystem in a Taylor series around the equilibrium geometry in Cartesian coordinates.

$$E_{e,n}(\mathbf{R}) = E_{e,n}(0) + \sum_A \left(\frac{\partial E_{e,n}}{\partial \mathbf{R}_A} \right) \mathbf{R}_A + \frac{1}{2} \sum_{AB} \left(\frac{\partial^2 E_{e,n}}{\partial \mathbf{R}_A \partial \mathbf{R}_B} \right) \mathbf{R}_A \mathbf{R}_B + \dots \tag{3.75}$$

The energy at equilibrium $E_{e,n}(0)$ can be set to zero. In addition, the first derivative at the equilibrium geometry is zero by definition. Hence, truncating the Taylor expansion after second order (harmonic approximation), the only remaining term is the second derivative matrix (Hessian). By a unitary transformation of the Hessian H

$$Q^T H Q = H_D \tag{3.76}$$

the Hessian can be diagonalized, thus in the basis Q a set of decoupled coordinates is obtained, i.e

$$\left(\frac{\partial^2 E_{e,n}}{\partial \mathbf{Q}_A \partial \mathbf{Q}_B} \right)_{A \neq B} = 0. \tag{3.77}$$

Since in this basis the nuclear kinetic energy is diagonal as well, a set of 3M independent differential equations can be obtained. All these equations have the form of the quantum mechanical harmonic oscillator whose solutions are known. In mass weighted coordinates, the angular frequencies of the vibrations are then obtained by the eigenvalues of the second derivative matrix. The Hessian of a system can be either computed by solving the coupled perturbed Kohn-Sham equations, or via numerical derivatives. Numerical derivatives are usually obtained with the three point central difference approximation. For a system with M particles this corresponds to 6M force evaluations.

3.5.2 Mode Selective Vibrational Analysis

For applications in which only few vibrations are of interest, mode selective vibrational analysis (MSVA)[23, 24] presents a possibility to reduce the computational cost significantly. Instead of computing the full second derivative matrix only a subspace is evaluated iteratively using the Davidson algorithm. As in NMA the starting point of this method is the eigenvalue equation

$$H^{(m)} q_k = \lambda_k q_k \tag{3.78}$$

with $H^{(m)}$ the mass weighted Hessian in Cartesian coordinates. Instead of numerically computing the Hessian elementwise, the action of the Hessian on a predefined or arbitrary displacement vector b

$$b = \sum_i b_i e_i^m \tag{3.79}$$

with the e_i^m i-th the mass weighted nuclear basis vectors, is computed. The action σ is given by

$$\sigma = H^{(m)} b = \begin{pmatrix} \frac{\partial^2 E_{e,n}}{\partial \mathbf{R}_1^{(m)} \partial b} \\ \frac{\partial^2 E_{e,n}}{\partial \mathbf{R}_2^{(m)} \partial b} \\ \vdots \\ \frac{\partial^2 E_{e,n}}{\partial \mathbf{R}_{3M}^{(m)} \partial b} \end{pmatrix}. \tag{3.80}$$

The first derivatives with respect to the nuclear positions is simply the force which can be computed analytically. The derivative of the force along components i with respect to b can then be evaluated as a numerical derivative using the three-point central difference approximation. The subspace approximation to the Hessian (Davidson matrix) in the i-th step is then obtained

as

$$H_{approx}^{(m),i} = B^T H^{(m)} B = B^T \Sigma. \tag{3.81}$$

The matrices B and Σ are a collection of the vectors b and σ computed up to the actual iteration step. $H_{approx}^{(m),i}$ is an i by i matrix, which, at convergence, includes the selected eigenvalue of the original Hessian. The approximate eigenvalues are obtained solving the eigenvalue problem

$$H_{approx}^{(m),i} u^i = d^i u^i \tag{3.82}$$

and from this the residuum vector (r^i) used as displacement vector for the next iteration step by calculating

$$r_m^i = \sum_k^i u_{m,k}^i \left[\sigma^k - d_m^i b^k\right] \tag{3.83}$$

The approximation to the normal coordinates q is finally obtained as

$$q \approx \sum_k^i u_{m,k}^i b^k \tag{3.84}$$

This method avoids the evaluation of complete Hessian, and therefore requires fewer force evaluations. It is important to note, that in the limit of 3M iterations the exact Hessian is obtained, thus the exact frequencies and normal modes in the limit of the numerical difference approximation.

3.5.3 Implementation

Tracked modes

The choice of the target mode in the iterative procedure can be done in several ways. So far in CP2K either a single mode, a range of frequencies or modes localized on preselected atoms can be chosen. In case, where a single mode is tracked, the residuum vector is calculated with respect to the mode with the eigenvalue closest to the input frequency. The method will only converge to the mode of interest if the initial guess is suitable. With the implemented algorithm always the mode closest to the input frequency is improved. Using an arbitrary initial guess, the mode of interest might not be present in the subspace at the beginning. It is important to note that the Davidson algorithm might converge before the desired mode becomes part of the subspace. Therefore, there is no warranty that the algorithm would always converges to the desired mode.

The second method allows to compute all modes in a range of frequencies. Here, the two cases have to be distinguished that either none or more than one mode is already present in the spectrum. In the first case, the mode closest to the desired range will be tracked. In the second case always the least converged mode will be improved. The advantage of this choice is that the major improvement with respect to convergence will be on this mode and in addition the expansion of the subspace slightly improves the other nonconverged modes as well. This way a faster convergence of all frequencies in the range is achieved. Similar to tracking a single frequency a bad initial guess might not cover the subspace, required to track all modes in the desired range.
The third possibility is a mode selected by its major contributions. Instead of basing the target vector on the eigenvalues, each step the approximations to the eigenvectors of the full Hessian are calculated. The relative contributions of the selected atoms to the norm of the eigenvector is then calculated, and the one with the largest is used for the next iteration step. This method can be extended to include the type of desired vibration as well. For this method the same problems as discussed above are valid.

Initial guess

From the discussion above it becomes clear, that the initial guess is crucial for the quality of the result. A trivial guess is a random displacement of either all or only specified atoms. This can be useful, if the characteristics of a mode to track is unknown, such as lattice vibrations, or a wide range of frequencies is tracked. For modes with known vibrations, an estimate of the eigenvectors can be used according to the involved atoms and characteristic displacements (symmetric stretch, asymmetric stretch).
A more general guess can be obtained by exploiting an approximate Hessian. NMA and MSVA require the molecule in an energetic minimum, thus a geometry optimization has to be performed before. Several algorithms like the Broyden Fletcher Goldfarb Shanno (BFGS) algorithm construct an approximate Hessian. From this Hessian, an initial guess can be obtained by diagonalizing it and taking the mode closest to the target. The quality of this guess strongly depends on the quality of the approximate Hessian. Finally, vibrational analysis using lower quality methods, like semi-empirical or force field methods, can be employed to obtain an initial guess for higher level methods.

Block Davidson algorithm

To efficiently use a parallel environment using the MSVA approach two strategies can be combined. The first one is parallelizing the force calculations required for the numerical derivatives. The second is using the block Davidson algorithm. In the latter approach the parallel environment is split into n subenvironments, each consisting out of m processors performing the force calculations. The initial vectors are constructed for all n environments, such that n orthonormal vectors are generated. After each iteration step, the new b and σ vectors are combined into a single B and Σ matrix. These matrices are then used to construct the approximate Hessian, and from this the n modes closest to the selected vectors are again distributed. As only few vectors have to be communicated between the processes in the block-Davidson algorithm, this method allows for an almost perfect parallel scaling on an arbitrary number of processors.

Chapter 4

Solvent Structure at the Active Interface

The acetonitrile / anatase(101) interface can be considered a prototypical interface between an oxide and a polar aprotic liquid, and is common in dye sensitized solar cells. Using first principles molecular dynamics simulations of a slab of TiO_2 in contact with neat acetonitrile (MeCN), the liquid structure near this interface has been characterized. Furthermore, in order to investigate properties that require extensive sampling, a classical force field to describe the MeCN/TiO_2 interaction has been optimized, and we show that this force field accurately describes the structure near the interface. We find a surprisingly strong interaction of MeCN with TiO_2, which leads to an ordered first MeCN layer displaying a significantly enhanced molecular dipole. The strong dipolar interactions between solvent molecules lead to pronounced layering further away from the interface, each successive layer having an alternate orientation of the molecular dipoles. At least seven distinct solvent layers (approximately 12 Angstrom) can be discerned in the orientational distribution function. The observed structure also strongly suppresses diffusion parallel to the interface in the first nanometer of the liquid. These results show that the properties of the liquid near the interface differ from those in the bulk, which suggests that solvation near the interface will be distinctly different from solvation in the bulk.

4.1 Computational setup and Validation

All molecular dynamics simulations reported in this work have been performed using the CP2K simulation package.[25] A time step of 0.5fs and the microcanonical ensemble (NVE) have been employed for both ab initio and classical MD simulations. Force field based molecular dynamics simulations are based on smooth particle mesh Ewald and neighborlist techniques, which are common in bio-molecular simulation packages and lead to efficient and linear scaling implementations. The derivation of the FF used, is discussed in detail in the following section.

The density functional implementation in CP2K (Quickstep) has recently been reviewed in detail in Ref. [26], and is based on the hybrid Gaussian plane wave (GPW) scheme.[27] In this scheme, an efficient and linear scaling algorithm for the calculation of the Kohn-Sham matrix is obtained through a dual representation of the electron density. Whereas the wavefunctions are always represented using Gaussians, the electron density can be represented either in Gaussians (the primary basis), or in plane waves (the auxiliary basis). The transformation of representation between these basis sets can be performed efficiently.[26] The advantage of the Gaussian basis set is its localized nature, which leads to sparse matrix representations of the required operators, and the fact that only a small number of basis functions (relative to e.g. a plane waves basis set) is required to represent the wavefunctions accurately. The advantage of using an auxiliary plane wave basis set for the density is that Fast Fourier techniques can be employed to compute the electrostatic (Hartree) energy, avoiding the traditional bottleneck of Gaussian based DFT codes. Wavefunction optimization is performed using the orbital transformation minimizer, which guarantees convergence and avoids diagonalization of the Hamiltonian matrix.[28] The Perdew-Burke-Ernzerhof (PBE) functional[29] and Goedecker-Teter-Hutter(GTH) pseudopotentials[30, 31, 32] have been employed for all DFT calculations. A 280Ry plane wave density cutoff has been applied. Carbon, Hydrogen and Nitrogen of the MeCN-molecules were described using a standard triple-ζ basis with two sets of polarization functions (TZV2P), which is known to yield satisfactory results for structure and dynamics of liquid acetonitrile.[33] For the oxide, a recently developed procedure to derive molecularly optimized basis sets[34] has been employed to derive basis sets for Titanium and Oxygen. Using 5 primitives, basis sets of double-ζ quality have been optimized on small Ti-O compounds. An important advantage of these molecularly optimized basis sets is that they can be diffuse, even in the condensed phase, without introducing near linear dependencies in the basis.[34] Near linear dependencies can cause instabilities in the SCF procedure, a problem encountered with more traditional split valence basis sets.

In order to assess the quality of our computational setup, we have compared the optimized geometry of a slab of TiO_2 containing 288 atoms in three layers between CP2K and CPMD[36]. The unit cell was $15.13 \times 20.45 \times 22.00$Å, including 11.5Å of vacuum to separate the slabs. Since CPMD and CP2K can use exactly the same functional, pseudopotentials, and geometries, this is a good test for the quality of the basis set. A 150Ry plane wave cutoff for the wavefunctions has been employed with CPMD. A comparison of both optimized structures showed only differences smaller then 10^{-3}Å. DFT studies by Lazzeri, Vittadini an Selloni[35] of different surface geometries provide

	length/Å	bond expansion	Ref. [35]
Ti1-O1	1.84	-8.19%	-8.60%
Ti1-O2	1.98	2.16%	2.00%
Ti1-O3	2.05	2.40%	3.30%
Ti1-O4	1.79	-7.83%	-8.40%
Ti2-O1	1.85	-4.94%	-5.00%
Ti2-O2	2.01	0.20%	0.30%
Ti2-O3	1.93	-0.46%	-0.30%
Ti2-O4	2.06	3.05%	5.10%
Ti2-O5	2.15	10.92%	8.70%
Ti3-O4	2.09	7.72%	5.20%
Ti3-O5	1.96	-2.10%	-1.20%
Ti3-O6	1.94	0.00%	0.00%

Table 4.1: Computed bond lengths and bond expansions for the relaxed surface of anatase(101). Results for the bond expansion are compared to the calculations in Ref. [35], The two non-equivalent bond lengths in the bulk are 2.002Å and 1.942Å. The labels of the atoms refer to the labels in Fig. 3 of Ref. [35].

a reference to judge the effect of slab size, in particular thickness (z direction) and the use of k-points. Selloni and co-workers have used 2 low-symmetry k points, and up to six layers of TiO_2. k-points are currently not available in CP2K, and instead larger unit cells have to be employed. For the simulations of the interface, this is not too problematic, as small unit cells would introduce a severe artificial periodicity for the liquid. As shown in Table 4.1, the surface relaxations computed for anatase (101) in this work are in good agreement with the reference values from Ref. [35] Similar results have been obtained for anatase(100) (not shown). The main difference in bond distances is for bonds away from the surface, which is due to the fact that the slabs used in our setup contain just three layers.

4.2 Derivation and Validation of the Force Fields

The parameters of the FF consist of three parts: 1) for the semi-conductor, 2) for the MeCN liquid, 3) for the interaction between the semi-conductor and MeCN molecules.

Parameters for the semi-conductor and for the liquid are available in

literature, parameters for the interaction between the two are derived in this work. The TiO_2 slab is described using point charges and a Buckingham potential as parameterized by Bandura and Kubicki[37]. We employ the same bulk potential for the bulk and the surface atoms. We have verified that this potential, originally tested on rutile, led to a stable structure for anatase, with reasonable agreement to the DFT structure. For the three layer slab of anatase(101) described above, the main difference between DFT and FF is a displacement of the atoms orthogonal to the surface, increasing its total thickness by about 0.3Å. The structure exposed to the liquid has a better accuracy. Along the [101] axis atoms match within less then 0.01Å, while displacements along the [010] vector are at most 0.05Å.

The MeCN-MeCN interactions are described using a potential based on point charges, and van der Waals interactions, while the molecular geometry is described using harmonic potentials for bonds and bends. The FF has been derived in Ref. [38] and has been shown to reproduce the structure of liquid MeCN as obtained with first principles simulations in Ref. [39].

	DFT	FF
single molecule	14.24	12.58
monolayer parallel	4.88	6.68
monolayer antiparallel	4.56	6.26

Table 4.2: Binding energies of MeCN to anatase(101) in kcal/mol/molecule as obtained with DFT and FF calculations. Binding energies are either for single molecule or for a monolayer of MeCN. In the monolayer, the molecules can be oriented in the same way (parallel, as shown in the left panel of Fig. 4.1), or alternate their orientation (antiparallel, as shown in the right panel of Fig. 4.1).

The FF for the TiO_2/MeCN interactions is derived here. The electrostatic interaction is through the point charges mentioned previously, and parameters for a Lennard-Jones potential are optimized matching binding energies and forces along a molecular dynamics trajectory as obtained from DFT calculations. To describe the interactions of MeCN with anatase a Lennard-Jones (LJ) potential is employed. Force matching was used to optimize the parameters of the LJ potential, trying at the same time to reproduce the binding energies for a set of configurations and the forces along a DFT trajectory. Three distinct binding geometries have been used to obtain reference binding energies on a three layer slab with a 15.13Å $\times$ 10.23Å $\times$ 25.00Å unit cell. One is a configuration of a single molecules interacting

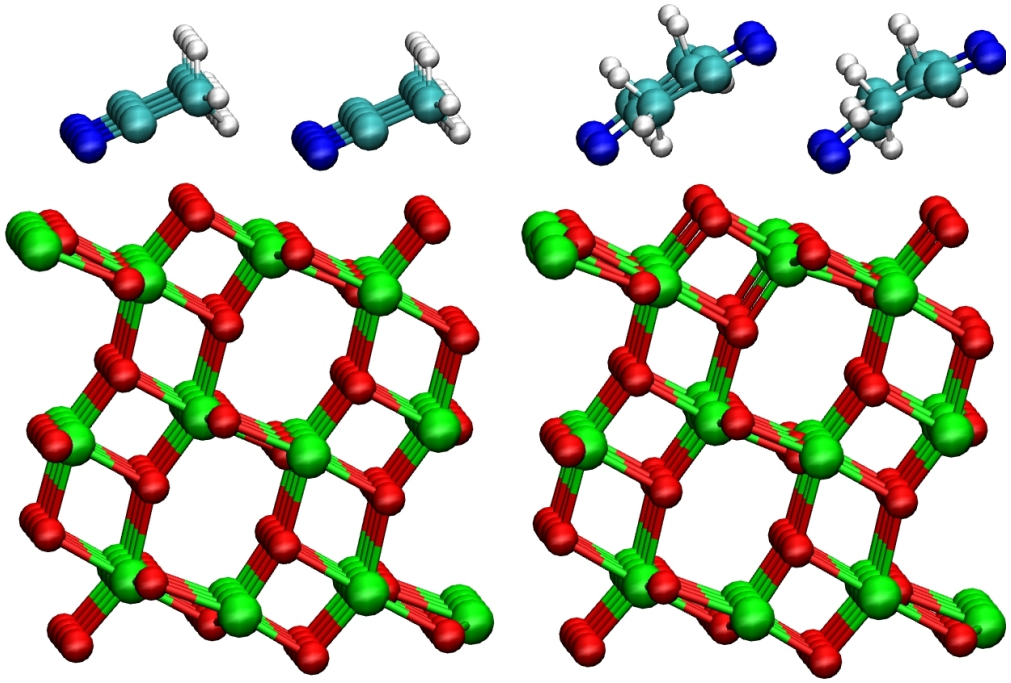

Figure 4.1: Shown are the optimized geometries of a monolayer of acetonitrile on a three layer anatase(101) slab. The left panel shows the parallel configuration, while the right panel shows the anti-parallel configuration. The parallel configuration is the lowest energy structure. However, the anti-parallel configuration is only 0.3 kcal/mol per molecule higher in energy. All atoms shown belong to the unit cell used, the three layer slab is fully optimized.

with the surface, in the other two cases a monolayer covered surface is used. We find that a single molecule binds with the carbonitrile group pointing towards the surface, i.e. with the negatively charged Nitrogen towards the positively charged Titanium. As shown in Table 4.2, the binding energy of this configuration is relatively large (14.24 kcal/mol). Configurations of a single molecule with the methyl group pointing towards the surface are not stable. Two different binding modes for the monolayer have been identified and are shown in Fig. 4.1. One with all molecules parallel to each other, and with an orientation as described for the single molecule, and one where the MeCN molecules alternate their orientation. As shown in Table 4.2, these two configurations have similar binding energies, the anti-parallel orienta-

tion being only slightly higher in energy. The binding energy per molecule is, however, significantly smaller than in the case of a single molecule binding the surface. The existence of an anti-parallel layer and the smaller binding energy per molecule are due to the intermolecular interactions. In the parallel case, MeCN molecular dipoles are aligned unfavorably with respect to each other, while in the anti-parallel case the molecular dipoles are oriented more favorably, but the interaction with the surface is unfavorable. Furthermore, favorable N–H interactions between molecules in different rows (with N–H distances of 2.52Å and 2.80Å) further stabilize the all-parallel structure. The all-parallel monolayer resembles to some extent the low-temperature (β) phase of solid acetonitrile[40]. From these monolayer calculations we can conclude that acetonitrile matches the structure of the anatase(101) surface very well.

The reference trajectory for the forces used in the force matching procedure was a short (800fs) first principles simulation with a three layer slab containing 216 atoms and 68 MeCN molecules of which 24 directly interacted with a surfaces of the slab. The unit cell is of this system is 22.69Å $\times$ 10.23Å $\times$ 39.00Å. This unit cell is such that it can accommodate the antiparallel monolayer, respecting its periodicity. In order to account for the two different possible binding geometries for acetonitrile, the molecules have been arranged so that half of them pointed with the carbonitrile group towards the surface and the other half with the methyl group.

i-j	ϵ_{ij},10^{-3} kcal*mol^{-1}	σ_{ij} , Å
H - O	6.144	2.905
H-Ti	2.064	2.613
C^c-O	2.397	4.423
C^c-Ti	5.905	2.673
N-O	2.922	4.303
N-Ti	0.701	4.131
C^m-O	5.833	1.681
C^m-Ti	7.263	4.134

Table 4.3: Lennard-Jones parameters for the MeCN/anatase interactions. C^c is the C atom of the carbonitrile group. C^m is the C atom of the methyl group

In order to obtain satisfactory results with the FF, the usual combination rules for the van der Waals parameters had to be abandoned, and instead ϵ and σ had to be defined for each pair of interactions. The parameters are

provided in Table 4.3. The particularly small value of ϵ implies that this part of the potential is nearly completely repulsive. The overall interaction energy, which is attractive, reproduces the DFT results in a satisfactory manner. For example, as shown in Table 4.2, for a single MeCN molecule binding to the surface, we obtain a binding energy of 12.58 kcal/mol with the FF, within 2 kcal/mol of the DFT results. Whereas a single molecule is underbinding,

		DFT	FF
single molecule	Ti-N	2.29	2.33
monolayer parallel	Ti-N	2.5	2.46
monolayer antiparallel^a	Ti-N	2.35	2.38
monolayer antiparallel^b	Ti-C^m	4.08	4.57
monolayer antiparallel^b	O-H	2.37	2.25

Table 4.4: Selected interaction distances for MeCN on anatase(101) in Å as obtained with DFT and FF calculations. For the antiparallel monolayer, the two non-equivalent molecules either point with (^{a}) the carbonitrile group or (^{b}) the methyl group towards the surface. C^m denotes the methyl Carbon.

the monolayers are overbinding by about 2 kcal/mol, but the relative stability of the parallel and anti-parallel monolayer is captured correctly. These results suggest that some cooperative effect at the interface is not fully captured by the FF model. Indeed, we will show in the following section that molecules near the surface exhibit a strongly enhanced molecular dipole, which is consistent with the above results. The quality of the obtained structure is illustrated in Table 4.4, where we report selected distances between atoms of the surface and of the substrate. For the strong Ti-N interaction, excellent agreement is found in all cases, and the deviation between the DFT and FF structure does not exceed 0.05Å. For those molecules pointing with the methyl group to the surface, slightly larger deviations are found, but for these configurations, the influence of the intermolecular MeCN interaction is important and these FF terms have been optimized for the neat liquid. As we will shown in the next section, the parallel orientation is dominant in an equilibrium simulation of the solid/liquid interface so that we consider the FF satisfactory.

4.3 Solvent structure near the interface

In this section, we investigate the structure of an MeCN liquid in contact with the anatase(101) surface. For the first principles studies, we have em-

ployed a unit cell of 22.69Å × 10.23 AA × 39.00Å, which contains 68 MeCN molecules and a 3 layer anatase slab. Both sides of the slab are in contact with the liquid, the region between the slabs is fully occupied by the liquid. This system contains more than 600 atoms, and approximately 2800 doubly occupied states are needed to describe the electronic structure. A 5ps trajectory has been computed. At the same time, FF based trajectories have been computed as well, one for exactly the same system (5ns), and one for a significantly larger system with a unit cell of 22.69Å × 20.46Å × 113Å (5ns). The first trajectory has been used to validate the classical model, while the larger model has been used to accurately compute properties. Already during the relatively short DFT trajectory, all sites on the the TiO_2 surface are occupied with MeCN, with most MeCN pointing with the carbonitrile group towards the corresponding Titanium atom. A second layer in the liquid can be seen, containing MeCN molecules that are typically antiparallel to the attached ones. In the classical simulation with the same cell, it can be seen that after 1ns all sites are occupied with MeCN , forming an all-parallel monolayer with defects. In the DFT simulation the average distance of a Titanium atom at the surface to the closest MeCN Nitrogen is 2.42Å, in the classical run this is 2.50Å. The standard deviation of the average values is in both runs in the range 0.16-0.18Å. A snapshot of the simulation is shown in Fig.4.2. It can clearly be seen that the first solvent layer resembles closely the all-parallel monolayer configuration. All molecules in the first layer adopt well defined locations in the plane of the surface. However, two kinds of orientational defects can be observed: molecules in an antiparallel configuration and in an orthogonal orientation. From our MD simulations, we can infer that the lifetime of these defects is the 100ps-1ns range. From this snapshot, also the second solvent layer can be discerned. In this layer, there is a tendency to form antiparallel dimers with molecules from this and other layers. Such antiparallel dimers are also found in the high-temperature (α) phase of solid acetonitrile.[40] The corrugation of the surface is visible in the first solvent and second second layer. To quantify these observations, we compute properties as a function of the distance to the surface, the z-position in the following. Shown in Fig. 4.3 are the Nitrogen density profiles as obtained with DFT and FF simulations. The DFT results, being averaged over only 5ps, exhibit more statistical noise than the FF simulations. Nevertheless, there is good agreement between these simulations, in particular concerning the peak positions. The difference in the height of the first peak can be attributed to the fact that the equilibrium concentration of the orientational defects has not yet been reached in the DFT simulation, while in the much longer FF simulation this is the case. The very pronounced first peak indicates a tightly bound Nitrogen atom, consistent with the calculations in the previous sec-

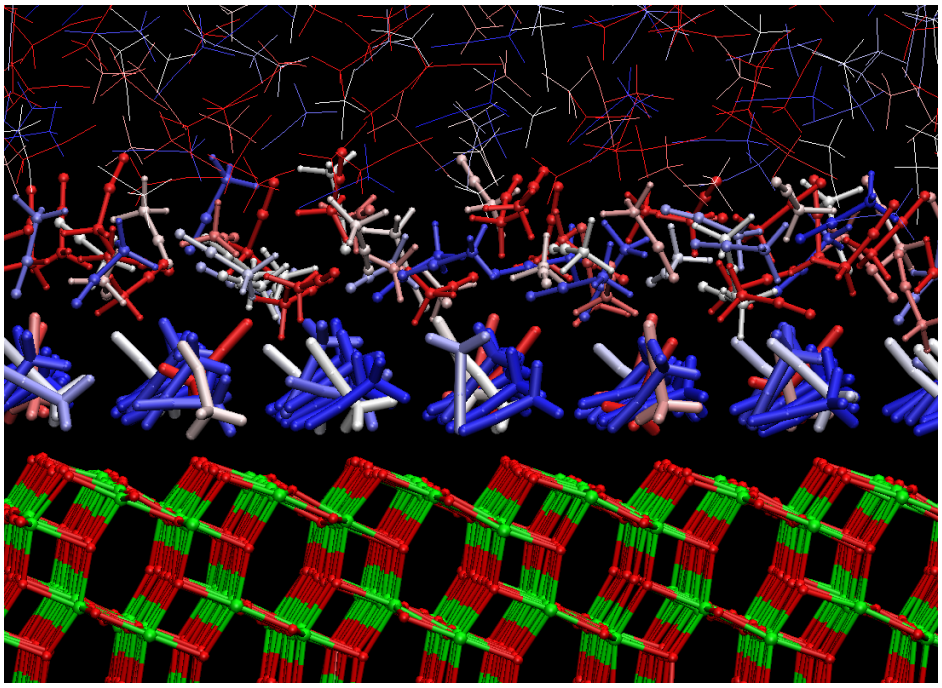

Figure 4.2: A snapshot from a classical simulation of the anatase(101)/acetonitrile interface. The solvent molecules are colored according to the cosine of the angle between their molecular axis and the 'ideal' axis as defined by the geometry of the all-parallel monolayer (see Fig. 4.1). The color scale goes from blue over white to red. Blue corresponds to a parallel, red to an antiparallel, and white to an orthogonal orientation. The first and second solvent layer (see text for details) are depicted using a tube and ball-and-stick representation respectively.

tion. Further layers, extending at least 10Å away from the surface, can be observed in the density profile, but in order to exclude sampling and size effects, it is useful to focus on the FF results obtained for the larger system, shown in Fig. 4.4. In this figure, several more peaks can be seen, at least up to 20Å away from the surface. The orientational defects seen in the first layer are the explanation for the appearance of an early second peak in the nitrogen profile at around 4.6Å, and the double peak (2.2Å and 3.5Å) for the central carbon. The alternating orientation of the solvent molecules can be most clearly illustrated with a figure showing the distribution of the cosines of the angle between the MeCN axis and the surface normal with respect to the distance between the surface and the center of mass of that molecules. Indeed, Fig. 4.5 displays at least 9 layers (17Å) with alternating orientation,

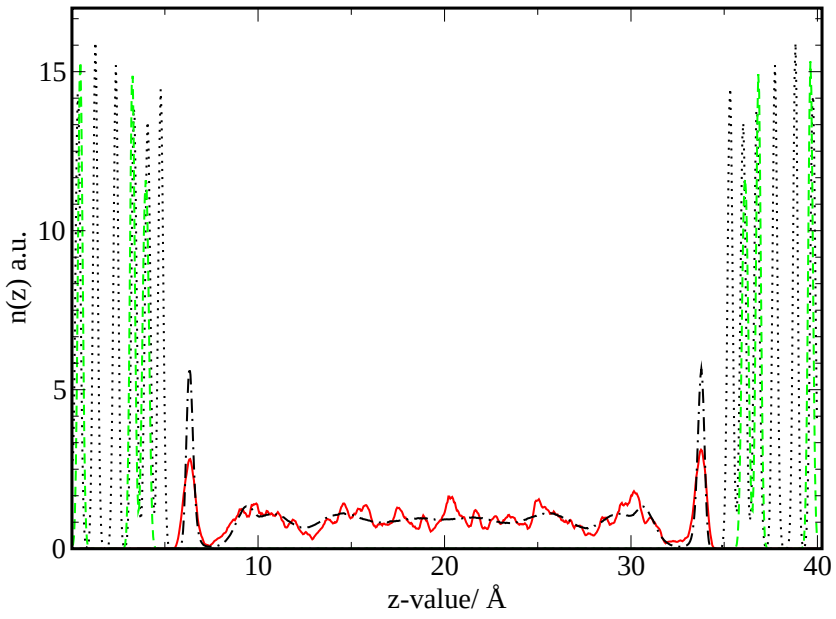

Figure 4.3: Shown are the density profiles of selected atom types for the DFT and FF simulations of the interface. The solid line (red) shows the Nitrogen distribution as obtained from the DFT simulations. This distribution can be compared to the same distribution as obtained from the FF simulations (dash-dotted line, black). The TiO_2 slab is represented by the FF distributions of the Ti (dashed line, green) and O (dotted line, black) as obtained from FF simulations.

and underlines the tendency of MeCN to pair its dipole. Note that each peak in the density profile corresponds to approximately two peaks in the orientational profile, suggesting that the upper and lower part of a layer (as defined by the density) have opposite orientations. The structure shown in Fig. 4.3 and Fig. 4.5 clearly suggests that the solvent properties near the interface will be significantly different from the properties in the bulk. The last property we investigate in this section are the molecular dipoles of the solvent molecules, since it can be expected that the TiO_2 surface will have a strongly polarizing effect on close by MeCN molecules. This property can not be obtained from a simple point charge FF model, but can be obtained from the DFT simulations using maximally localized Wannier functions[41, 42]. We have evaluated the molecular dipoles every 12.5fs during the ab initio molecular dynamics simulation. We find a significant increase of the dipole from about 5.0D in the liquid to approximately 6.3D for those molecules directly interacting with the surface. However, the shift in the average dipole

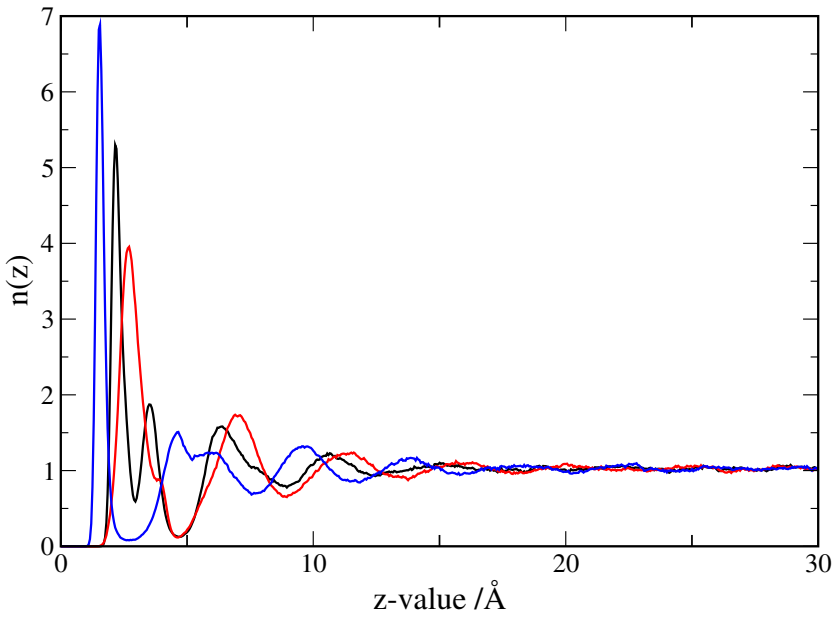

Figure 4.4: Shown are the density profiles of the liquid phase near the interface as obtained for the FF simulations of the interface using the large simulation cell (113Å z-edge). The first three peaks, from left to right, are Nitrogen (blue), central Carbon (black), and methyl Carbon (red). The zero of the x-axis corresponds to the position of the first oxygen layer.

is already much reduced (5.2D) in the second layer. The strongly enhanced dipole of the molecules in the first layer can in part explain the difficulty experienced by the FF to reproduce exactly the detailed energetics of the interactions as shown in Table 4.2. We note however, that a partial charge transfer from the Nitrogen atom to the Titanium atom, i.e. from the solvent to the solid, would appear in this analysis as an enhanced molecular dipole. Furthermore, such a charge transfer might be artificially large due to the self interaction error present in local approximations to density functional theory. The distance dependent distributions of the molecular dipoles (not shown) indicate that the fluctuations of the dipoles around the average value are up to five times larger in the first two layers than in the bulk.

4.4 Solvent dynamics near the interface

In this section, the influence of the TiO_2 surface on the self diffusion of MeCN will investigated, and the diffusion constant as a function of the distance with

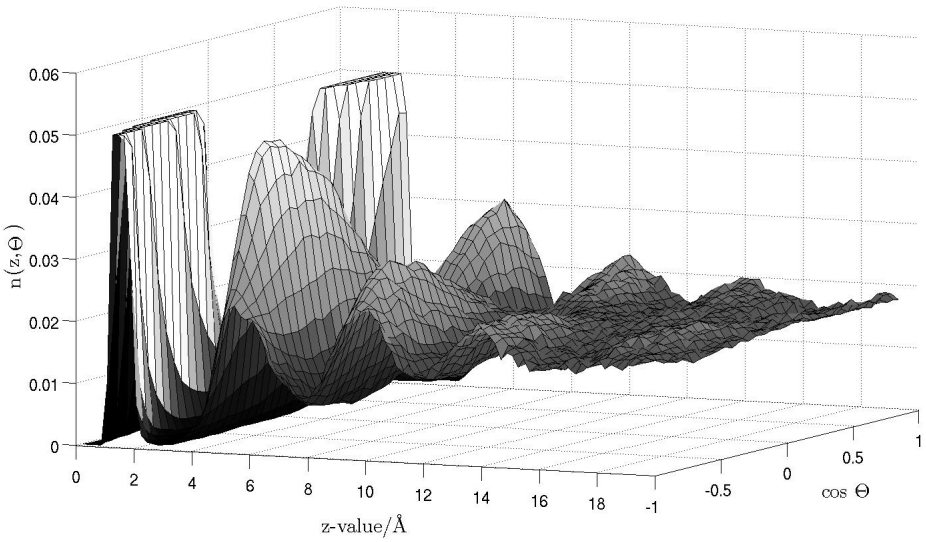

Figure 4.5: The distribution of the cosine of the angle between the surface normal and the MeCN dipole is shown with respect to the distance between the center of mass of the molecule and the surface. The alternating orientation of the successive solvent layer can clearly be seen.

respect to the surface will be computed. These calculations are based on the approach of Liu, Harder and Berne [43] for calculating anisotropic diffusion constants near an interface. The central idea is to calculate the mean square displacements for those molecules i staying in the layer $[a, b]$ during the entire period $[t, t + \tau]$:

$$\langle \Delta x^2(\tau) \rangle_{[a,b]} = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{N(t,t)} \sum_{i \in [a,b]_{[t,t+\tau]}} (x_i(t+\tau) - x_i(t))^2 \qquad (4.1)$$

This function is then weighted by the probability for an particle to stay the time τ in this layer:

$$P(\tau) = \frac{1}{T} \sum_{t=1}^{T} \frac{N(t, t+\tau)}{N(t,t)}, \qquad (4.2)$$

where $N(t, t+\tau)$ denotes the number of particles that remained in the interval $[a, b]$ during the entire period $[t, t + \tau]$. So that the self diffusion constant for a given layer can be calculated as:

$$D_{xx}([a,b]) = \lim_{t \to \infty} \frac{\langle \Delta x^2(\tau) \rangle_{[a,b]}}{2\tau P(\tau)} \qquad (4.3)$$

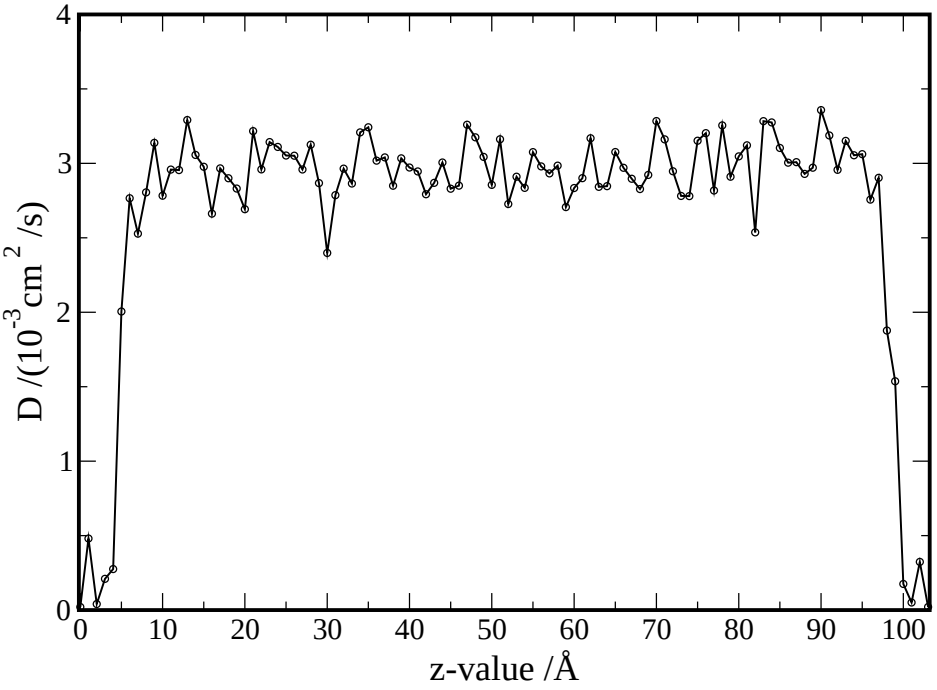

Figure 4.6: Shown is the self diffusion constant of MeCN for diffusion parallel to the surface. The strongly suppressed diffusion near the interface can clearly be seen, despite the statistical noise due to the procedure used for computing the distance dependent diffusion constant (see text for details).

These formulas have been applied to the larger classical system, intervals [a,b] of 1Å lead to the best statistical accuracy. Nevertheless, as the function $P(\tau)$ decays relatively quickly to zero, only a short part of the mean square displacement curve can be employed to obtain an estimate of the diffusion constant, and some statistical uncertainty can not be avoided. The results are shown in Fig. 4.6. Far away from the interface, the diffusion constant is fluctuating around 3x10^{-3} cm^2/s. This in good agreement with the diffusion constant we have calculated for the bulk acetonitrile experimental density using the formula for diffusion in an isotropic liquid (2.93x10^{-3} cm^2/s). This indicates that our simulation of the interface is properly equilibrated and at the correct pressure. Near the interface, at a distance of about 10Å, the diffusion constant starts decreasing till it reaches a value close to zero in the first layer of MeCN. This behavior can be explained by the molecular structure we have described in the previous section. Indeed, the relatively strong binding of the MeCN molecules to the TiO_2 surface immobilizes the first layer. The strong dipolar interactions between the subsequent layers of

alternating orientation and the corrugation of the surface introduces a barrier for diffusion. We have observed that in these layers diffusion is a collective defect mechanism in which one molecules rotates out of the layer leaving a vacancy. This in turn allows for diffusion in the layer.

4.5 Discussion

Solids have traditionally been the domain of the plane wave community, while molecules have traditionally been studied with atom centered basis sets such as Gaussians. In this work, we have for the first time investigated a solid-liquid interface using the CP2K simulation package and the GPW approach. Our results indicate that this approach is well suited for both the molecular liquid and the slab, even though we find that the traditional split valence basis sets that dominate the molecular calculations are not well suited for calculations on solids. The recently proposed molecularly optimized basis sets seem to perform well. The advantage of the GPW approach is that large systems, containing several hundreds of atoms, can be simulated efficiently, even allowing for Born-Oppenheimer molecular dynamics. Both capabilities are required for simulating solid/liquid interfaces, as such systems must be large and must be treated at a finite temperature, i.e. in the regime where the liquid exists as such. Nevertheless, despite the fact that a first DFT simulation of 1ns has been performed using CP2K[44], long timescales (>1ns) generally remain inaccessible to first principles molecular dynamics simulations. It is therefore required to employ computationally less demanding methods if these timescales must be probed. In this work, we have also used CP2K to perform FF based simulations of the interface. Not only is it convenient to be able to switch quickly between simulation methods, also the flexibility of the classical code, e.g. mixing potential types suited for various systems within one simulation, is advantageous for these systems. The force field has been derived and validated using first principles results, with satisfactory results for interaction energies and geometries. Nevertheless, we are aware of the limitations of a non-polarizable force field for the description of this system, and we have emphasized the strong polarization of molecules near the interface.

The interface we have studied, MeCN in contact with anatase(101), shows a rich structure. We find that the solvent interacts rather strongly with the solid, forming a well defined first solvent layer that might passivate the surface in DSSCs. The strong dipolar interactions between molecules and the fixed orientation of the molecules in the first layer introduce a pronounced layering of the solvent near the interface, with a tendency to form layers of

alternating orientation. This structure extends, depending on the criterion used, up to 10 - 20Å away from the interface. For DSSCs this is the region which contains the dye, and where light induced electron transfer and the dye re-reduction take place. From the FF simulations, we find a much reduced diffusion in the interfacial layer, which could imply that mass transport in the narrow channels of the mesoporous TiO_2 could be lower than anticipated. Furthermore, since the solvent structure near the interface differs significantly from the bulk structure, we expect solvent properties, e.g. the ability to solvate ions or the solvent dielectric constant to be different for the interface region and the bulk liquid.

Chapter 5

Structure of the Dye Semiconductor Interface

In dye sensitized solar cells, three structurally similar dyes are commonly employed to sensitize anatase nano-crystals, namely the cis-bis(isothiocyanato) bis(2,2'-bipyridyl-4,4'-dicarboxylato)-ruthenium(II) dye (N3), its twice deprotonated (N719) and completely deprotonated (N712) form. Using density functional theory, several possible binding geometries of these dyes are identified on the anatase(101) surface. Computed relative energies show that protonation of the surface can strongly influence the relative stabilities of these configurations, and could induce a conformational transition from double bidentate-bridged binding to mixed bidentate/monodentate binding. Attenuated total reflection (ATR)-IR experiments and computed vibrational spectra provide additional support for a protonation dependent equilibrium between two different configurations. Furthermore, self-assembly in chains of hydrogen bonded dye molecules seems structurally favorable on the anatase(101) surface, for enantiopure dyes a packing density of $0.744/\mathrm{nm}^2$ could be achieved.

5.1 Structures and Energetics

It is intuitive that dyes like the N3 dye use their carboxylate groups to bind the TiO_2 surface. On the top layer of the anatase(101) surface, the most favorable anchoring points for these groups are the undercoordinated titanium atoms (5–fold, Ti^{5c}). However, several surface sites and four carboxylate groups are available for binding, and three different coordination types for a carboxylate group to a metal oxide – monodentate, bidentate bridging and bidentate chelating – are known. Therefore a wide range of binding geometries is in principle possible. Here, based on systematic computations of the relative energies of these configurations, a limited subset of likely configurations is identified.

The carboxylate groups can be either trans or cis to a SCN-ligand, so even with a single anchoring group, two different orientations of the dye are possible. Computed relative stabilities show that singly anchored dyes are less

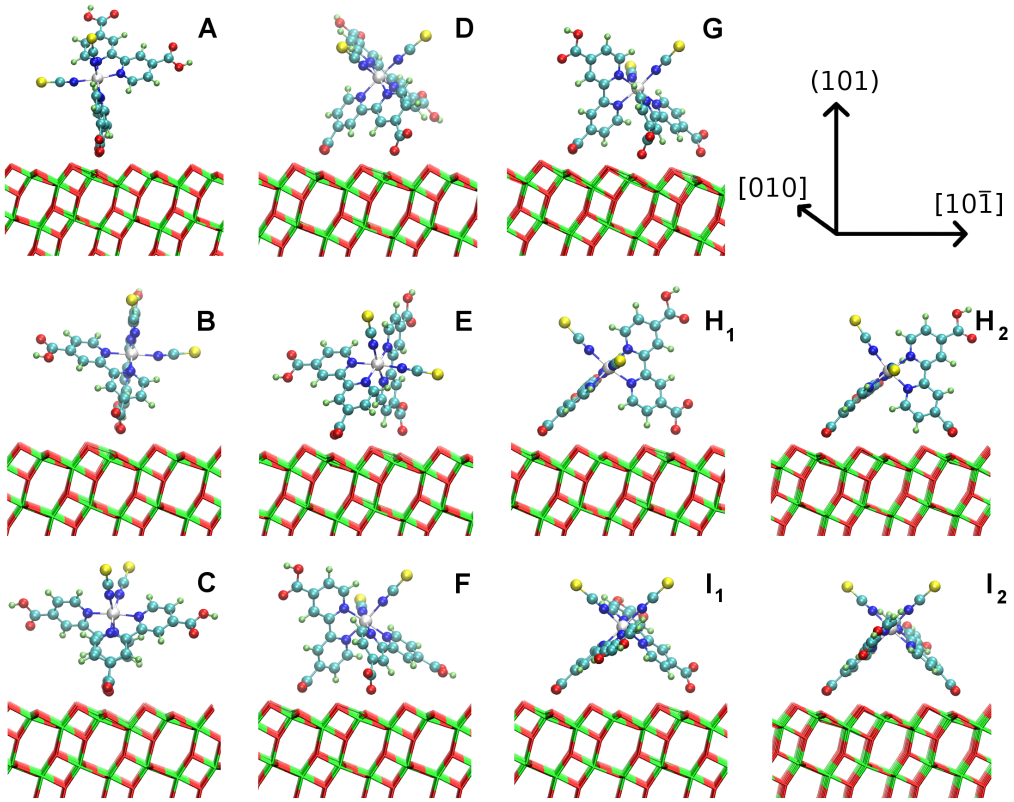

Figure 5.1: The different possible orientations and coordination patterns of the stable two–fold and three–fold anchored dyes bound to the anatase (101) surface are shown. Differences can be either found in the spacing along the [10$\bar{1}$]-vector, the relative position of the anchoring carboxylate groups or the coordination type of the the carboxylate group to the surface. A top view of the $\mathbf{I}_2$, in a dense packing arrangement, is shown in Fig. 5.7.

stable than dyes that have two or three groups binding to the surface. This is consistent with spectroscopic measurements indicating that two groups are involved in the anchoring of the dye[45, 46]. None of optimized structures binds in a bidentate chelating way. We therefore focus on the structures shown in Fig. 5.1 that all feature two or three binding carboxylate groups either in a bidentate bridging or monodentate coordination. First, we discuss the case in which all binding groups are available in a deprotonated state, and no surface protons are present.

The double anchored structures can be divided into three different classes depending on the distance between the anchoring groups along the [10$\bar{1}$] vector of the surface. Structures **A**, **B** and **C** are bound to Ti^{5c} without spacing

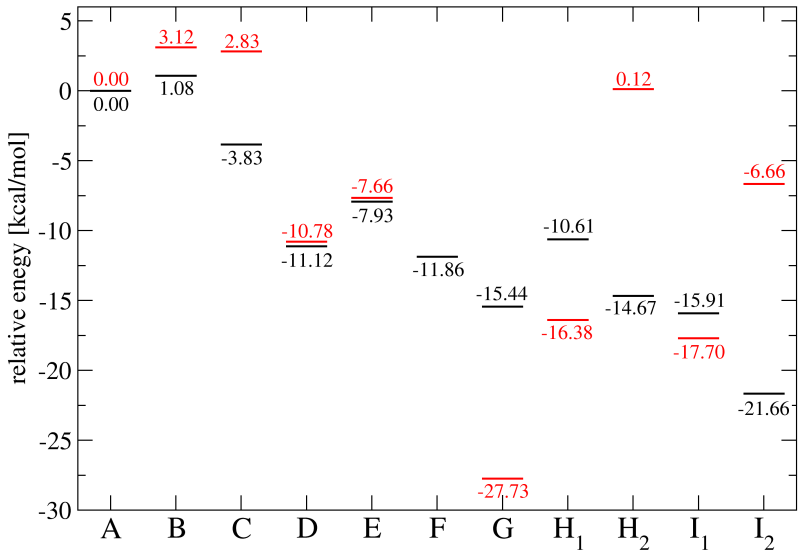

Figure 5.2: Shown are the relative stabilities with respect to **A** of the different dye conformations shown with their label in Fig 5.1. All energies as computed on a 3 layer slab. Black data refers to the proton–free surface, while red data refers to the situation where protons are added to the surface surface, and interacting favorably with monodentate bound carboxylate groups. This strong interaction can shift the relative stability of two configurations, for example for $\mathbf{I}_1/\mathbf{I}_2$ and $\mathbf{H}_1/\mathbf{H}_2$.

along this vector, i.e. bound to the Ti sites along the [010] vector. As shown in Fig. 5.2, these three structures are approximately 20 kcal/mol higher in energy than the most stable configurations found. Hence, they are unlikely to occur in the actual device. Their low stability can be explained by the distortion of the dye structure that is required to match the carboxylate group position to the Ti^{5c}. Comparing the coordination types of the three configurations, an interesting effect can be observed. In **A** the carboxylate group trans to the SCN group is coordinating bidentate bridging while the cis group prefers the monodentate coordination. The same preference of coordination type is found for **B** and **C**, where either both groups coordinate monodentate or bidentate according to their relative position to the SCN ligand. Since in all three configurations the environment of the carboxylate groups is equivalent, the explanation for this behavior could be the trans

influence of the SCN group. Also in the more stable configurations discussed below, this effect is likely to persist.
D, **E** and **F** present the possible structures binding to neighboring Ti^{5c} along the $[10\bar{1}]$ vector. Energetically these structures are of intermediate stability but still about 7-10kcal higher in energy than the most stable configuration. In all three structures a mixed monodentate/bridging coordination pattern is found. Here, interactions of the biisonicotinic acid ligand with the surface, in particular the top layer of undercoordinated oxygen (O^{2c}), are sufficient to explain these differences. Essentially, if there is no steric interaction between the biisonicotinic acid ligand and the O^{2c}, a bidentate bridging coordination type is preferred. A monodentate coordination is adopted if strong interactions between the O^{2c} and the ligand would result from bridged binding. Indeed, the monodentate coordination allows for a larger spacing between the ligand and the surface, and the energetic penalty of adopting a less stable monodentate coordination, might be balanced by the reduced steric interaction. For **D**, **E** and **F**, only the mixed bidentate/monodentate coordination was found to be stable. From **F** it is possible to derive a structure bound via three carboxylate groups to the surface (**G**). This structure is the only stable triple–bound structure we have identified. However, the energy gained by going from a double bound to a triple bound structure is modest. In the actual cell, the triple bound systems has to displace more solvent molecules from the surface. As estimated in Ref. [47], the interaction energy between acetonitrile and the anatase(101) surface can reach 4-5 kcal/mol for a single acetonitrile molecule interacting with a neat surface. In presence of a solvent, the total energy balance might thus be in favor of the double bound configurations.
The most stable double anchored structures, **H** and **I**, are obtained for maximal spacing between the carboxylate groups along the $[10\bar{1}]$ vector. Both the mixed bidentate/monodentate binding ($\mathbf{H}_1$ and $\mathbf{I}_1$) and the double bidentate-bridged binding ($\mathbf{H}_2$ and $\mathbf{I}_2$) are stable. Note that even in the $\mathbf{I}_2$ the two ligands are not equivalent, and only one of the two ligands can adopt the monodentate configuration to yield the mixed binding $\mathbf{I}_1$. The double bidentate configurations are lower in energy than the mixed configurations. In agreement with earlier suggestions[48], $\mathbf{I}_2$ is the most stable configuration. Typical Ti–O(dye) bond lengths are 2.08Å for the bidentate-bridged bonds, and 2.02Å for the monodentate bonds.
The discussion above referred exclusively to relative energies computed for dyes interacting with a defect-free, clean surface. Solvent and surface defects might influence the relative stabilities of the various configurations. As shown in Fig. 5.2, in the presence of surface protons the relative stabilities of the configurations are strongly influenced. For this setup, a reorientation of **F**

into **G** was found, as the hydrogen bond at the monodentate group induced a small rotation of the dye, allowing for stronger interactions of the free carboxylate group with the surface. Monodentate binding carboxylate groups can form a strong 'hydrogen bond' to a nearby surface bound proton, while bridged bidentate groups can not. In the absence of the solvent, the stabilization due to this interaction is approximately 17 kcal/mol. This is reasonable if one considers that the carboxylate group involved in this bond is a charged group, and that this hydrogen bond thus has the character of a salt bridge. In the presence of (surface) protons, the double bridged configurations ($\mathbf{H}_2$ and $\mathbf{I}_2$) are thus relatively unfavorable, and mixed bridged/monodentate forms (**G**, $\mathbf{H}_1$ and $\mathbf{I}_1$) are energetically preferred. In the presence of a solvent, the stabilization due to hydrogen bonding is expected to be smaller than in the gas phase[49], especially for the conformation **G** which forms a low barrier hydrogen bond. Nevertheless, it is clear that protonation can influence the binding geometry. Since interconversion between $\mathbf{H}_1$ and $\mathbf{H}_2$, and $\mathbf{I}_1$ and $\mathbf{I}_2$ is facile, a pH-dependent equilibrium between these configurations is to be expected. We present further theoretical and experimental evidence for such an equilibrium in the next section.

5.2 IR-experiments and computations

Using mode selective vibrational analysis [23], frequencies in the range from 1100cm^{-1} to 2200cm^{-1} have been computed for all dye geometries **A**–**H** as optimized on the neat and protonated surface. As shown in Fig. 5.3, a good agreement between calculations and experiment is obtained. However, a vibrational analysis for these large condensed phase systems is computationally demanding and technically challenging, and some differences remain. In particular, changing the total charge of the simulated system, for example changing from a neat to a protonated surface, can induce a shift of 30 cm^{-1} in sensitive modes such as the stretching vibrations of the carboxyl groups. Similarly, the computed intensity is likely to be strongly influenced by the restriction of vibrational analysis to solvent free systems. This is due to the fact that the intensity is related to the derivative of the system dipole, which is poorly screened in the absence of solvent. Furthermore, for all solvent free calculations, free carboxyl groups are left protonated to reduce the overall charge of the system, while experimentally they can also be deprotonated. A detailed one to one comparison of computed and experimental spectra is thus not straightforward, but calculations can assist the assignments of the experimental spectra. Consistent with our estimated precision,

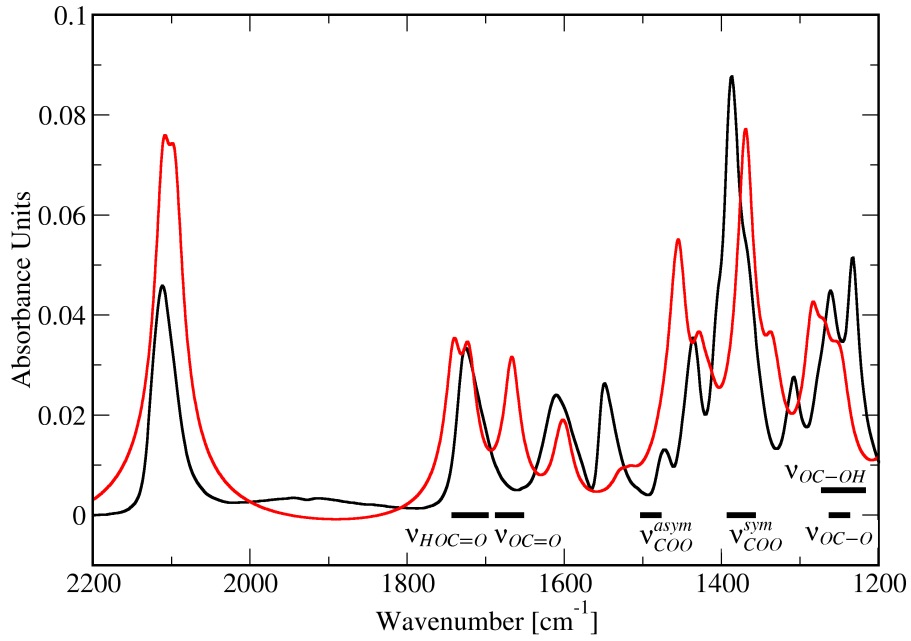

Figure 5.3: Shown are the experimental IR spectrum of the N3 dye adsorbed on anatase nano-crystals in EtOH(black) and the computed IR spectrum (red) for the fully protonated configuration $\mathbf{I}_1$ on the anatase(101) surface. As discussed in the text, most spectral features match to within the accuracy of the computational setup.

we therefore only report an average frequency, rounded to five wavenumbers, for each coordination type. Except for **B**, all configurations have at least one carboxylate group coordinating bidentate bridging to the surface. The ν_{COO}^{asym} for this groups is found at 1480-1500 cm^{-1} in all configurations. The ν_{COO}^{sym} is somewhat more sensitive to the orientation of the dye, but appears always in a range between 1360 to 1390 cm^{-1}. For the monodentate groups in configuration **D**,**E**,**F**,$\mathbf{I}_1$,$\mathbf{H}_1$ the $\nu_{OC=O}$ is found close to 1670 cm^{-1} and, with slightly larger spread, the ν_{OC-O} between 1240 and 1260 cm^{-1}. Finally, the free carboxyl groups give signals at 1000-1100 cm^{-1} (δ_{OCO-H}), 1220-1270 cm^{-1}(ν_{OC-OH}) and 1700-1740 cm^{-1} ($\nu_{HOC=O}$). The computed frequencies are in good agreement with signals found in experiment (Fig. 5.3 and Fig. 5.4), but can not be used to uniquely identify a particular binding geometry. The difficulty in identifying a particular configuration is in part

due to the presence of unspecific signals such as ring modes in the experimental spectra that overlap with potentially revealing signals such as ν_{COO}^{asym}. At the same time, computed normal modes are often fairly complex as a significant mixing between elementary vibrations with similar frequencies is observed (see supplementary materials). Note that these overlaps might render these ring modes unsuitable for the purpose of normalization, as their intensity might be influenced by the binding geometry.

The experimental ATR-IR measurements have been performed on the N3 (sequence 1), N712 (sequence 2) and N719 (sequence 3 and 4) dyes adsorbed on sintered anatase nano-particles. These dyes are initially protonated, deprotonated and partially protonated respectively. In each sequence, the pro-

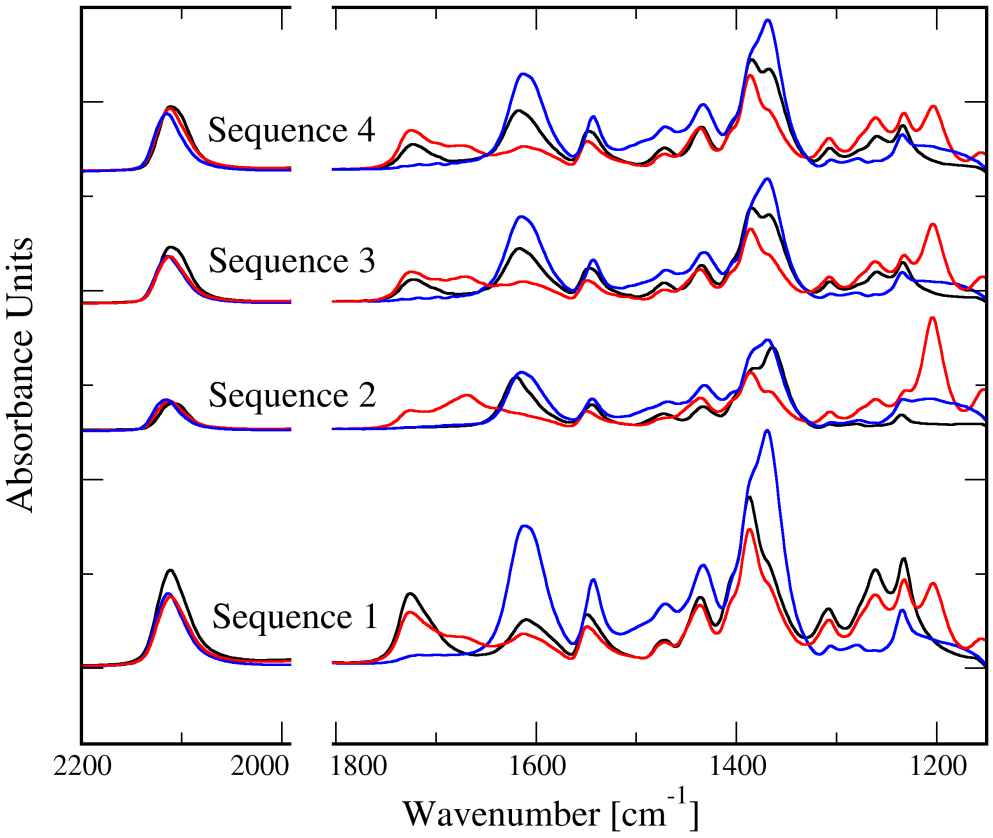

Figure 5.4: Shown are the experimental IR spectra for the N3 (sequence 1), N712 (sequence 2), N719 (sequence 3 and 4) dye on mesoporous TiO_2 while flowing either neat ethanol (black), or solutions of ammonia (blue) and TFA (red). For all spectra, the neat solvent in contact with TiO_2 has been used as a background, and TFA and ammonia signals thus remain in the spectrum.

tonation has been varied after adsorption by treatment with the acid TFA and the base ammonia, which provides detailed data on the spectral features related to protonation and deprotonation of the system. The adsorption process of each dye was monitored by ATR spectroscopy, ending with the corresponding black spectrum shown in Fig. 5.4. In sequence 1 and 3, the proton containing N3 and N719 dyes were then contacted with a solution of dissolved ammonia (blue) and finally with dissolved TFA (red). In sequence 2 and 4, a similar experiment is performed with the N712 and N719 dyes. Here, the adsorbed dyes were contacted first with TFA (red) and then ammonia (blue). In sequence 1, the $\nu_{HOC=O}$ signal at 1728cm^{-1} completely vanishes upon treatment with ammonia, indicating that base is strong enough to deprotonate the dye. Comparing the integrated intensities of the SCN signals (2112 cm^{-1}) in sequence 2, this signal is found to be almost independent of the protonation state of the dye. This suggests that, starting from the dye in a deprotonated state, ammonia does not desorb the dye, and that protonation does not influence the SCN intensity. The latter observation is in contrast to the suggestion in Ref. [46], where the difference in intensity between N3 and N712 was attributed to a difference in induced dipole moment of the SCN group. Note that, if the dye is adsorbed in the presence of protons (e.g. sequence 1), subsequent ammonia treatment results in a partial desorption. Since the intensity of the SCN peak is less for the N712 dye than for the N3 or N719 dyes, and since the intensity is protonation independent, this indicates that less dye is adsorbed if N712 is employed.

To highlight the observed differences shown above, we present in Fig. 5.5 the computed spectra of $\mathbf{I}_1$ and $\mathbf{I}_2$ and the experimental difference spectrum of the N3 and N712 dyes adsorbed on mesoporous TiO_2 in the neat solvent. Before subtraction, the experimental spectra have been normalized to the intensity of the SCN peak. Qualitatively similar spectra can be obtained subtracting the spectra for the N712 dye in contact with ammonia and TFA, avoiding this normalization procedure. However, in that case additional TFA and ammonia signals lead to a more complicated spectrum. In the difference spectra, the fitted Lorentzians in green account for more intense signals in a protic environment, while the Lorentzians in blue indicate stronger signals in the absence of protons. Comparing the experimental difference spectrum to the computed spectra of $\mathbf{I}_1$ and $\mathbf{I}_2$, the more intense $\mathbf{I}_1$ resp. $\mathbf{I}_2$ signals are generally found at the location of the fitted green, resp. blue Lorentzians. As previously described, the computed signals at 1500 and 1670 cm^{-1}, which correspond to the ν^{sym}_{COO} and ν_{OC-O} vibrations, only match the experimental frequencies if the estimated error of 30 cm^{-1} is taken into account. The difference in signals at 1720 and 1380 cm^{-1} are related to the protonation state of the free carboxylate groups, and thus not included in the computed

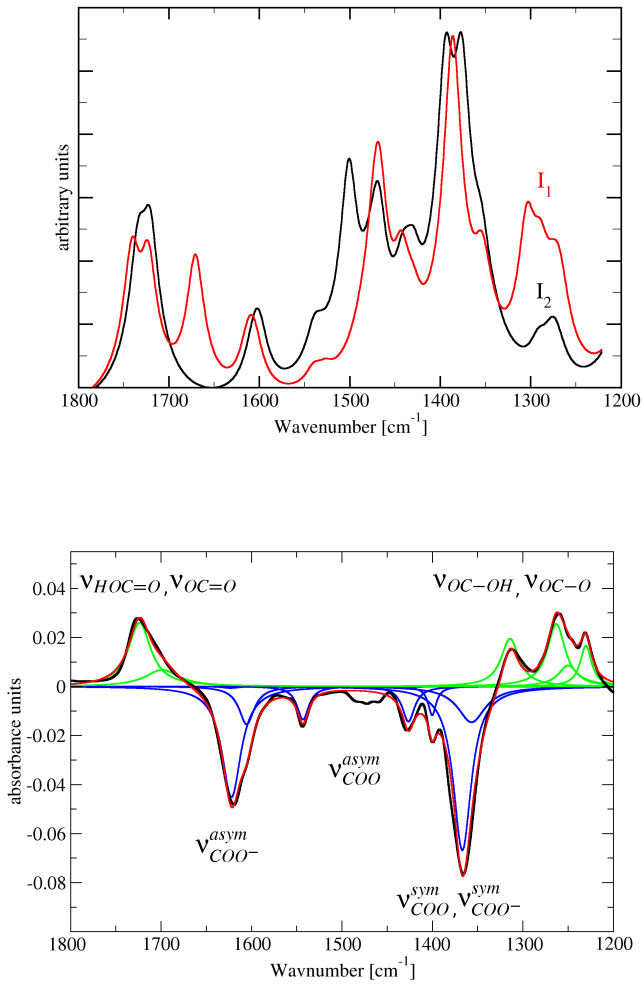

Figure 5.5: On top, the computed spectra for configurations $\mathbf{I}_1$ (red) and $\mathbf{I}_2$ (black) are shown. On the bottom, the experimental difference spectrum (black line) of the neat N3 and N712 dyes adsorbed on mesoporous TiO_2 is shown. This spectrum has been fitted (red line), the individual Lorentzians in green and blue indicate more intense signals in the N3 and N712 spectrum respectively. Assignments based on calculations are shown near the measured bands, but are not intended to resolve individual peaks. It can be observed that more intense signals in the computed $\mathbf{I}_1$ and $\mathbf{I}_2$ spectra generally correspond to green (N3) and blue (N712) Lorentzians respectively.

spectrum. Taking these effects into account, the computed spectra matches experiment rather well, and provides further evidence for a pH dependent equilibrium of $\mathbf{I}_1$ and $\mathbf{I}_2$.

5.3 Discussion

Our calculations of the dye on the anatase(101) surface suggest that configurations **H** and **I**, which are both favorable from energetic point of view, can interconvert from a double bidentate-bridged configuration to a configuration with mixed bidentate/monodentate binding. While the double bidentate-bridged configuration is the more stable one in the absence of protons, the single monodentate configuration can be stabilized strongly in the presence of protons. The experimental spectrum for the N3 dye shows increased intensity for regions that are expected in monodentate configurations, but these could not be assigned uniquely, since all of them superimpose with unspecific signals like ν_{OC-OH} and $\nu_{HOC=O}$. For the N712 dye, the vibrational difference spectrum displays an increased intensity for peaks that can be assigned to bridged binding, including a combined asymmetric stretch mode of the two carboxylate groups, which could be an indication of a double bidentate-bridged configuration. Further support for a protonation dependent binding geometry comes from ab initio molecular dynamics simulations of the full interface, i.e. oxide, dye and solvent. These simulations have been started with configuration $\mathbf{I}_2$ in the absence of protons, and ran for 8ps without change in binding mode. However, adding two protons to the surface, but not directly interacting with the dye, led to a spontaneous change from $\mathbf{I}_2$ to $\mathbf{I}_1$ after 1.5 ps of simulation. Also the latter configuration, shown in Fig. 5.6, was stable for a further 8ps of simulation. These results provide further theoretical evidence that both states are stable in solution, and that interconversion could be rapid. It also suggests that not only direct hydrogen bonding to the dye determines the binding energy, but that there might be an indirect effect through changes in the electronic structure of the surface due to protonation. Experimentally, the initial protonation state of the dye is known to influence the solar cell efficiency.[46] We therefore suggest that protons might affect the system's efficiency not only by changing the relative position of the electronic levels of dye and oxide, but also by changing the binding geometry. For example, the biisonicotinic acid ligands are less equivalent in configuration $\mathbf{I}_1$ than in the configuration $\mathbf{I}_2$, especially in the proton stabilized configuration. The electronic coupling between dye and oxide is thus likely to be different in these configurations. A quantification of this effect is beyond the scope of the current investigations.

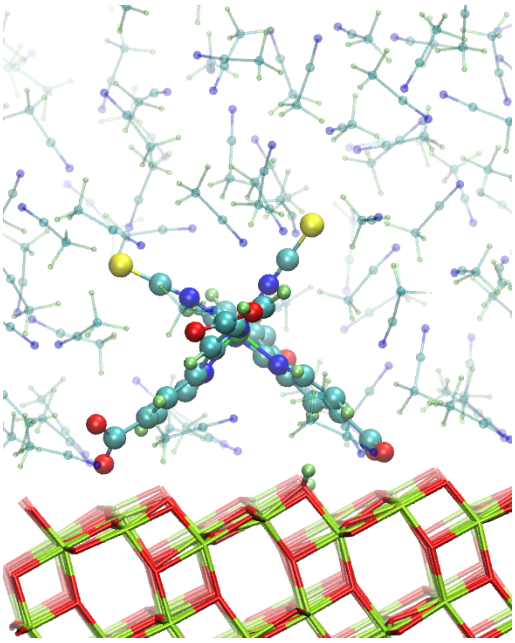

Figure 5.6: Shown is a snapshot of an ab initio molecular dynamics simulations in which oxide, dye and solvent have been treated explicitly. In this simulation, a transition from the $\mathbf{I}_2$ to the shown $\mathbf{I}_1$ configuration has been observed after addition of surface bound protons. This snapshot furthermore illustrates that the ordering of the first acetonitrile layer[47] persists in the neighborhood of the adsorbed dye. In this configuration, both SCN groups are exposed to the solvent, and the sulfur atoms are approximately 10Å above the interface.

Finally, we would like to discuss the observed partial desorption of the N3 dye when the system is being treated with ammonia, while this desorption is not observed for the N712 dye. Even though the different behavior of the two dyes could possibly be related to a difference in binding geometry (e.g. $\mathbf{I}_1$ vs. $\mathbf{I}_2$), we believe that the interconversion between these two states is too rapid to explain the difference in desorption. Instead, we propose that the N3 and N719 dyes can form stable multimers on the surface that interact favorably with neighbors through the unbound carboxylate groups. Short chains and clusters of the N3 dye at low coverage on a rutile surface have been observed in STM experiments[50], and several dense packing arrangements have been discussed in Ref. [48]. Shown in Fig. 5.7, is a densely packed arrangement of linear chains that share a single proton between two carboxylate groups. Geometry optimization of this system yields hydrogen bonds that are 2.38Å long, a favorable distance for a strong interaction. Note that only a single

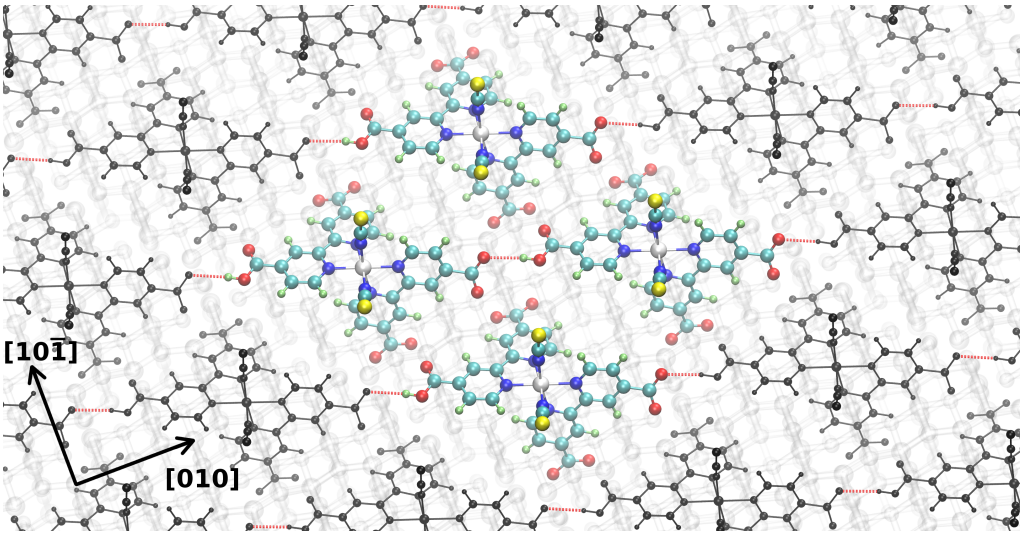

Figure 5.7: Shown is the optimized structure of the most dense packing of configuration $\mathbf{I}_2$ on the anatase(101) surface. In this configuration, each dye binds the oxide with the two carboxylate groups trans to the SCN ligands, while the other two carboxylate groups interact with neighboring dyes through hydrogen bonds. The bond length of the hydrogen bonds, drawn in red, is 2.38Å, the surface coverage of the dye in this assembly is 0.744/nm^2. The basic unit cell contains only one dye, for the geometry optimization a larger super cell (shown for clarity only with color for the dyes) has been employed.

proton per dye molecule is required for this packing. Deprotonation of the system with ammonia would leave the system in an energetically unfavorable state, with charged carboxylate groups in close proximity. In a densely packed arrangement, desorption could result. The N712 dye, in the absence of protons, cannot form these chains and thus cannot reach the same dense packing. As a result less dye will adsorb initially and ammonia treatment should have no effect, in agreement with our experimental observations. Note that two different enantiomers of the N3-like dyes exist (Λ-N3 and Δ-N3). Due to the surface symmetry, chirality is not important for the interactions of a single dye with the anatase(101) surface. However, the self-assembly presented in Fig. 5.7 is only possible for enantiopure dyes. A racemic mixture of molecules might not display long range order, or, as has been observed experimentally for heptahelicene on Cu(111)[51], form smaller domains of ordered dyes, or self-assemble with a different pattern. We therefore suggest that the use of enantiopure dyes could lead to a denser dye packing, which in turn could lead to improved device characteristics. We observe that in

the self-assembled monolayer, all four carboxylate groups have a structural role, either binding to the surface, or to neighboring dyes. At the same time, the SCN groups are well exposed to the solvent, which will facilitate dye regeneration through S-I^- interactions and complex formation [52]. As formation of hydrogen bonded chains requires a certain mobility of the dyes on the surface, the mechanism of diffusion of the dye on the surface is of interest. Diffusion of carboxylate bound molecules on TiO_2 surfaces has been discussed in literature, and even for multiple-anchored but flexible molecules barriers as low as 14 kcal/mol have been reported.[53, 54] This would suggest that diffusion and thus self-assembly is possible at the temperatures (20-80C) and timescales (hours) employed experimentally for adsorbing the dye. Here, we have investigated one mechanism for surface diffusion along the $[10\bar{1}]$ direction, assuming that diffusion could proceed as a sequence of rotations of the COO group around the C–C bond, going from bidentate via a monodentate transition state to bidentate at a new surface cite. We obtained an approximate value of 40 kcal/mol for the energy barrier of this mechanism by performing constrained geometry optimizations of the dye on the neat surface, starting from the $\mathbf{I}_2$ configuration and rotating both binding carboxylate groups simultaneously. If a non-concerted mechanism would be possible, despite the relatively rigid shape of the dye, this large barrier could be significantly reduced. Alternatively, the solvent or available protons could stabilize the transition state strongly by hydrogen bonding. Identifying in an accurate and systematic way the mechanism for diffusion is beyond the scope of the current work.

5.4 Methods

All density functional theory calculations were performed using the CP2K program package[25]. DFT calculations are based on the hybrid Gaussian and plane wave (GPW) scheme with the Perdew-Burke-Enzerhof[29] exchange correlation functional and corresponding pseudo potentials[31, 32]. The solvent in the molecular dynamics simulations has been described with Gaussian basis sets of double ζ quality with one polarization function, while a more accurate triple-ζ basis with double polarization has been employed for the dye[26]. The exception is ruthenium, for which highly accurate molecularly optimized [55] basis sets have been used. Basis sets for Titanium and surface oxygens have been optimized separately the same way as described [55]. For Titanium a DFT+U correction has been applied, such that the computed bandgap matches the experimental, to partially compensate for self interaction errors. The plane wave density cutoff was set to 400 Ry. The periodically

repeated unit cell measures $22.69 \times 20.45 Å^2$, and a three layer slab has been used for all calculations (144 TiO_2 units). To verify the impact of the slab thickness, we have reoptimized the $\mathbf{I}_1$ and $\mathbf{I}_2$ configurations (protonated) on a six layer slab, the relative stability was 11 and 7 kcal/mol on the three and six layer slab respectively. This difference is non-negligible, but might be similar in magnitude to the effect of other approximations in the computational setup. The reduced energy difference for the larger model would add further support for the proposed equilibrium between configurations. Geometry optimizations have been performed using the Broyden-Fletcher-Goldfarb-Shanno algorithm using a maximum displacement of 0.0002Å as convergence criterium, and allowing all atoms to relax. Vibrational analysis of the dye on the surface was performed employing mode selective vibrational analysis [23] as implemented in CP2K converging all modes in a range between $1100 cm^{-1}$ to 2200 cm^{-1}. The initial configuration for ab initio molecular dynamics have been obtained by 1ns preequilibration of the solvent, keeping the other position fixed, using force field based molecular dynamics. The FF for acetonitrile is based on point charges, van der Waals interactions and harmonic bonded terms as found in literature[38] , while the interaction between the surface and solvent has been derived and validated against ab initio data in our earlier work[47], and solvent-dye interactions the dye has been parametrised based on the generalized Amber force field. The full system was afterwards equilibrated for another 6 ps using ab initio MD, in the NVE ensemble, with the settings described above, but a reduced cutoff of 280Ry.

Infrared spectra were measured on a Bruker IFS 66/S FT-IR spectrometer equipped with a dedicated ATR-IR attachment (Optispec) and a liquid nitrogen cooled MCT detector. All spectra were recorded at a resolution of $4 cm^{-1}$. The titania layered Ge prism was fixed within an in house built stainless steel flow cell. The volume of the cell (< 0.052 ml) was defined by a viton O-ring fitting into a groove in the steel cell body. The flow-through cell was cooled by means of a thermostat and the measurements were performed at 40℃. The flow of liquid over the sample was controlled by means of a peristaltic pump (ISMATEC Reglo 100) located behind the cell. Liquid was provided from three separate glass reservoirs. The flow from the reservoirs was determined by a computer controlled pneumatically actuated three way Teflon valve (Parker PV-1-2324).

Deposition of titania (Ti-Nanoxide D, Solaronix CH) onto the Ge (Harrick, 50 x 20 x 1 mm) internal reflection element (IRE) was performed as described by Solaronix (http://www.solaronix.ch/technology/assembly/). Prior to the film preparation, ca. 10 nm of TiO_2 was deposited on the Ge IRE by physical vapor deposition (PVD). This was necessary to fix the Nanoxide particles

on the IRE. Fixation was achieved by calcining the dried Nanoxide film on the prepared Ge IRE at 460℃ for 1h. Dye adsorption on the PVD titania showed that the ATR-IR signals were negligibly small compared to the signals obtained on the Nanoxide films. The dyes (N3, N719 and N712) were purchased from Solaronix. Ethanol (EtOH, Fluka 99% puriss) was used as received. Trifluoracetic acid (TFA, Fluka) and ammonia (2M solution in ethanol, Aldrich) were dissolved in EtOH to obtain corresponding solutions of 2mM. Assuming refractive indices of 2.5 for the wet film and 4.0 for Ge, a penetration depth dp of 1.2 μm (0.4 μm) at 1000 cm^{-1} (3000 cm^{-1}) was calculated.

5.5 Conclusions

By a combination of in-situ ATR-IR experiments and computations the binding behavior of the N3, N712 and N719 dyes on anatase(101) surfaces has been studied. Energetics from solvent free geometry optimizations show that the spacing of the binding carboxylate groups along the $[10\bar{1}]$ vector is an important factor for the stability of the different configurations, and that both biisonicotinic acids ligands are involved in binding. In the absence of solvent molecules, three configurations **G**, **H** and **I** are significantly more stable than the other configurations. A strong influence of surface protonation and hydrogen bonds on the relative stabilities of the various configurations has been observed, favoring a double bidentate bridging binding ($\mathbf{H}_2$ and $\mathbf{I}_2$) type in the absence of protons, and a mixed bidentate/monodentate binding in the presence of protons ($\mathbf{H}_1$ and $\mathbf{I}_1$). A comparison of the experimental IR data and computed vibrational frequencies supports that an equilibrium between these states is present at the interface. MD simulations of the interface with explicit solvent and the **I** configuration suggest that interconversion could rapid. Provided a proton source is available, as for the N3 and N719 dyes, the energetically favorable configurations $\mathbf{I}_1$ and $\mathbf{I}_2$ can form chains of dye molecules on the surface. This further stabilizes these configurations, and allows for a densely packed layer of dye molecules. Long range order will only be possible for enantiopure dyes.

Chapter 6

Sensitizer Regeneration and Ion Distribution

A highly efficient mechanism for the regeneration of the cis-bis(isothiocyanato) bis(2,2'-bipyridyl-4,4'-dicarboxylato)-ruthenium(II) sensitizing dye (N3) by I^- in acetonitrile has been identified using molecular dynamics simulation based on density functional theory. Barrier–free complex formation of the oxidized dye with both I^- and I_2^-, and facile dissociation of I_2^- and I_3^- from the reduced dye are key steps in this process. In situ vibrational spectroscopy confirms the reversible binding of I_2 to the thiocyanate group. Additionally, simulations of the electrolyte near the interface suggest that acetonitrile is able to cover the (101) surface of anatase with a passivating layer that inhibits direct contact of the redox mediator with the oxide, and that the solvent structure specifically enhances the concentration of I^- at a distance which further favors rapid dye regeneration.

6.1 Regeneration of the dye

In an effort to elucidate the regeneration mechanism of the dye, previous spectroscopic studies have identified a short-lived I_2^--radical intermediate[56]. Additionally, spectroscopic evidence for the existence of an intermediate dye-iodine complex has been presented[57, 56, 58], and recent DFT studies on the isolated molecule verified that iodide can indeed bind to the oxidized dye[59]. Here, based on simulations in the condensed phase that probe not only iodide but also the I_2^- and the formation of I_3^-, we propose a complete regeneration mechanism, which we partially verify with experiments. The concentration of I_2^- is much lower than that of I^-, but I_2^- plays a crucial role in the full mechanism. We qualitatively confirm the earlier calculations, which predict binding of the iodide ion and I_2^- radical to the dye.[59] Most importantly, we extend these studies to explicitly include the solvent, so that we can provide thermodynamical data, and insight in the association/dissociation kinetics. In the gas phase, both I^- and I_2^- can bind the oxidized dye either covalently at the sulfur atom of the SCN group, or more weakly via bipyridyl iodide

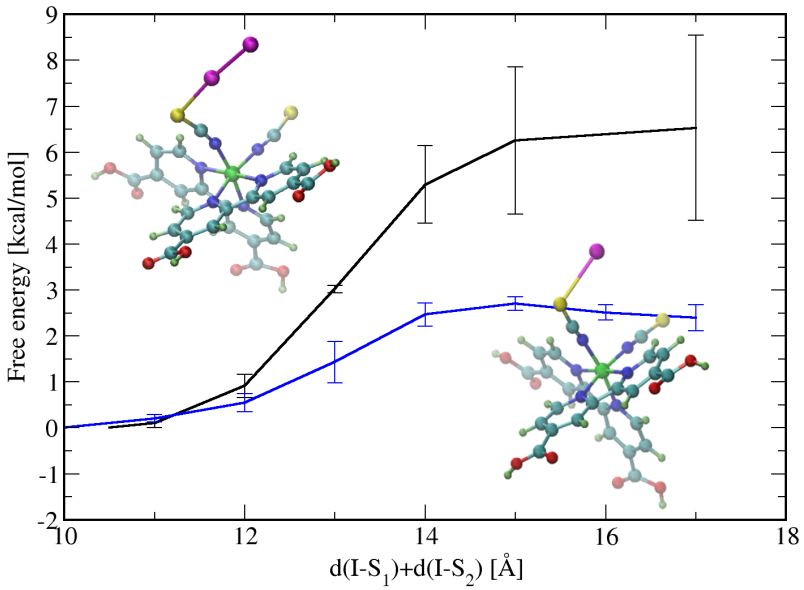

Figure 6.1: Shown are the free energy profiles for association of the [dye:I] complex (lower curve) and [dye:I_2] complex (upper curve) in acetonitrile solution. The reaction coordinate employed (x-axis) is the sum of the two S-I distances (see text for details). Both complexes are thus weakly stable, and association is barrierless. The inserts show snapshots of the stable complexes in solution, the solvent is not shown.

interactions. Interestingly, in the first case, the central metal is reduced, the complex thus involves a neutral I(0) or I_2, while the charge is not transferred for the weakly bound complex. However, simulations of the solvated complexes show that only the sulfur bound complexes are stable in acetonitrile, while the bipyridyl bound structures dissociate within 1 ps of simulation without electron transfer to the dye. Consequently, we focus on the stable sulfur bound [dye(II):I(0)] and [dye(II):I_2] complexes in the remainder of this work. To quantify the stability of these complexes, and to get insight in their association kinetics, we have computed the free energy profiles of dissociation using thermodynamic integration along a specified reaction coordinate. For this calculation, it is useful to take as a reaction coordinate the sum of the distances between iodide and the sulfur atoms of both SCN groups. This guarantees that dissociation is a smooth process, i.e. the recombination

of the ion with the dye is prevented as the reaction coordinate increases. These results, shown in Fig. 6.1, confirm the stability of both complexes, and predict a binding free energy of about 3–4 kcal/mol and 5–6kcal/mol for the [dye,I] and [dye,I_2] complex respectively. Additionally, we find that the free energy profile is barrierless, which indicates that the reaction rate for association will be diffusion limited. As shown in Fig. 6.2, we have been

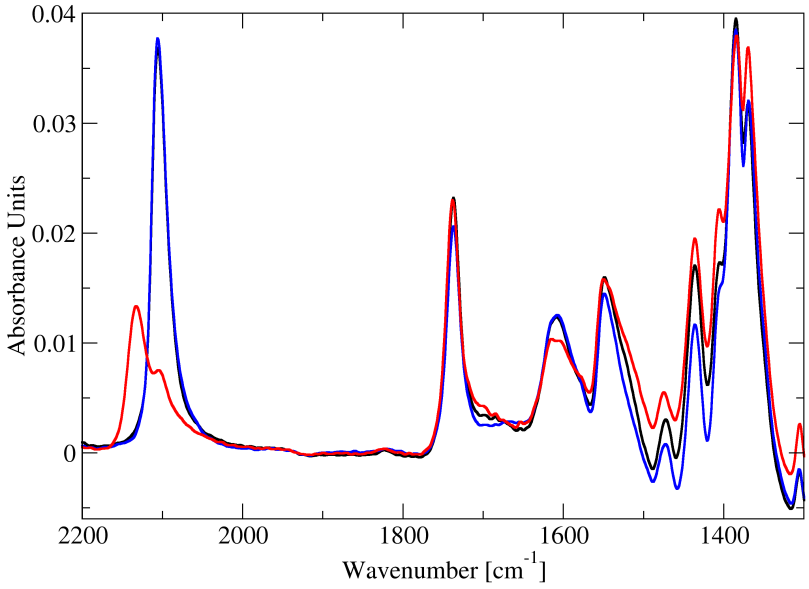

Figure 6.2: Experimental IR-spectra of the N3 dye attached to TiO_2 in acetonitrile, before adding iodine (black line), with I_2 added (red line), and after washing iodine away (blue line). The pronounced peaks at 2098cm^{-1} and 2130cm^{-1} correspond to vibrations of the SCN and I_2SCN ligands, and support our finding of reversible complex formation.

able to detect this complex in situ by a sequence of ATR-IR experiments on the reduced N3 dye attached to anatase nano–particles in contact with acetonitrile and molecular I_2. The spectrum of the dye attached to the surface shows a single high intensity peak at 2098cm^{-1} corresponding to the C-N stretching mode of the SCN-ligand. After the addition of I_2, the peak splits into two signals, one remaining at 2098cm^{-1} with lower intensity, and a new one at 2130cm^{-1}. The reduced intensity of the SCN-ligand signal, together with the new signal at 2130cm^{-1}, can be explained by the formation of an

I_2SCN ligand. The assignment of the peaks was verified by computing the gas phase spectrum of the [dye,I_2] complex, in which we found the signal of the I_2SCN ligand to be about 40cm^{-1} higher than for the SCN-ligand. The computed weak binding free energies are in good agreement with the experimental observation that the binding of I_2 is easily reversible, i.e. flowing acetonitrile through the cell restores the original signal at 2098cm^{-1} and eliminates the peak at 2130cm^{-1}. In order to investigate the remainder of

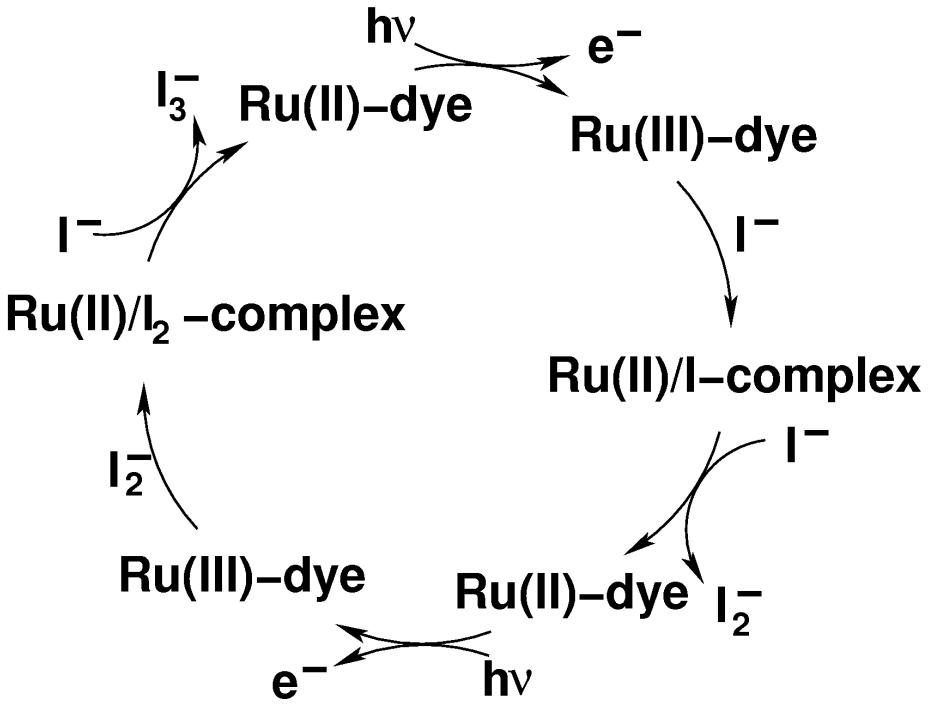

Figure 6.3: Proposed mechanism of the dye regeneration. In a barrier free way, two oxidized dyes are reduced by I^-, and I_3^- is produced.

the regeneration process, we have added additional I^- and I_2^- ions to our simulations of the [dye,I] and [dye,I_2] complexes. An interaction of I^- with the [dye,I] complex has been suggested before.[57, 56] In all cases, this led to an immediate dissociation of the complexes, as soon as the additional iodide reached the binding side of the halide, yielding I_2^- or I_3^- and a reduced dye. These results combined suggest an efficient pathway for dye regeneration under working conditions in DSSC, which is shown in Fig. 6.3. The first part of the reaction is the interaction of I^- with the oxidized dye, yielding a stably bound [dye:I] complex. This complex interacts with I^- to form a reduced dye and free I_2^-. Contrary to the assumption that I_3^- is formed in the

electrolyte by disproportionation of I_2^-, we propose that I_2^- instead interacts with another oxidized dye and forms a [dye:I_2] complex or that I_2^- interacts with a [dye:I] complex to directly form I_3^-. Additionally, also the [dye:I_2] complex can further interact with I^- to yield I_3^-. This mechanism has the advantage that it does not require two I_2^- radicals, which are only present in low concentration, to meet in bulk solution. Instead, the radical anion can react in place of formation with a photo-oxidized dye close by. Remarkably, the full pathway proceeds in a barrier free way. This is in agreement with the fast regeneration rates observed in experiments and the unrivaled efficiency of this combination of dye and redox pair.

6.2 Ion distribution at the interface

The proposed formation of various dye–iodide complexes during dye regeneration raises the question of the availability of iodide in this active region. Furthermore, the influence of the solvent, and in particular the excellent performance of acetonitrile, has not been fully explained yet. To address this question, we employ atomistic models of the interface based on the (101) surface of anatase, various ions, and acetonitrile. First, we compute the probability of finding an ion at a given distance from the surface using thermodynamic integration (TI). The probability profiles can directly be interpreted as concentration profiles, and are shown in Fig. 6.4 for a number of monovalent anions (F^-, Cl^- and I^-), cations (Li^+, Na^+, K^+), and I_2^-. All curves exhibit a non-monotonic behavior, with several distinct minima and maxima, contrary to the smooth distribution that a Poisson-Boltzmann description of this system would predict. Calculations of the distance dependent solvent dielectric constant, shown in Figure 6.5 , indeed suggest that local linear response is not any longer valid in this region and that the solvent at the interface can not be described with a single dielectric constant.[60, 61] This indicates the limitations of continuum theory in this important region, and suggests that explicit atomistic models might also enhance theoretical predictions of interfacial electron transfer based on continuum models.[62] Interestingly, the concentration of I^- strongly peaks at about 10Å from the surface, and exceeds the bulk concentration, while the concentration of all other mono-atomic ions studied decreases towards the surface. A high concentration of I^- near the surface, in particular near the SCN groups, which can be expected to be at about 10Å, will contribute to a rapid reduction of oxidized dyes, enhancing the performance of DSSCs. In the case of the I_2^- radical, the structure is less pronounced, but there is certainly no strong thermodynamic gradient that would favor diffusion away from the interface.

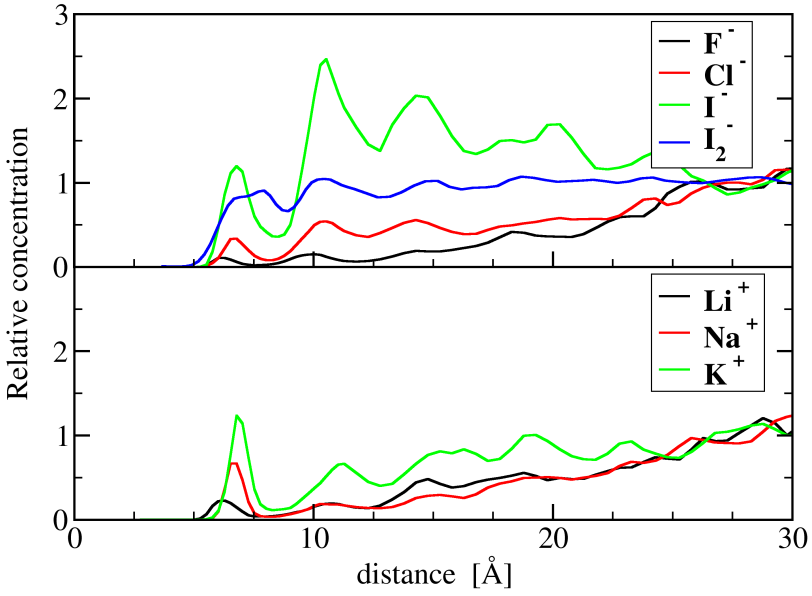

Figure 6.4: The concentration, relative to the bulk value, of various monovalent ions and I_2^- as a function of the distance to the anatase (101) surface. Data obtained from thermodynamic integration in the neat, otherwise salt-free solvent. I^- is present in higher concentration near the interface, while the other monovalent ions are being repelled from the interface.

This supports our proposed reaction mechanism, which suggests that I_2^- further reacts with the [dye:I] complex to produce I_3^- near the interface. We note that bromide has previously been used in X-ray reflectivity studies of the counter ion distribution near a lipid monolayer[63]. In that study, a minor second maximum could be discerned in the ion distribution, but the experimental resolution did not allow quantification of the small feature. We believe that the pronounced effect we observe in this system, for a heavy ion near a potentially flat surface, could be verified experimentally. The ionic profiles discussed so far have been obtained by simulations that mimic very low ionic strength. The observed structure can therefore not be attributed to a correlation between ions, but must be induced by the solvent. In previous chapter, we have observed a strong layering of the acetonitrile near the anatase(101) surface, due to a favorable interaction of the nitrile group with the titanium atoms of the surface, and the pronounced dipolar nature of the

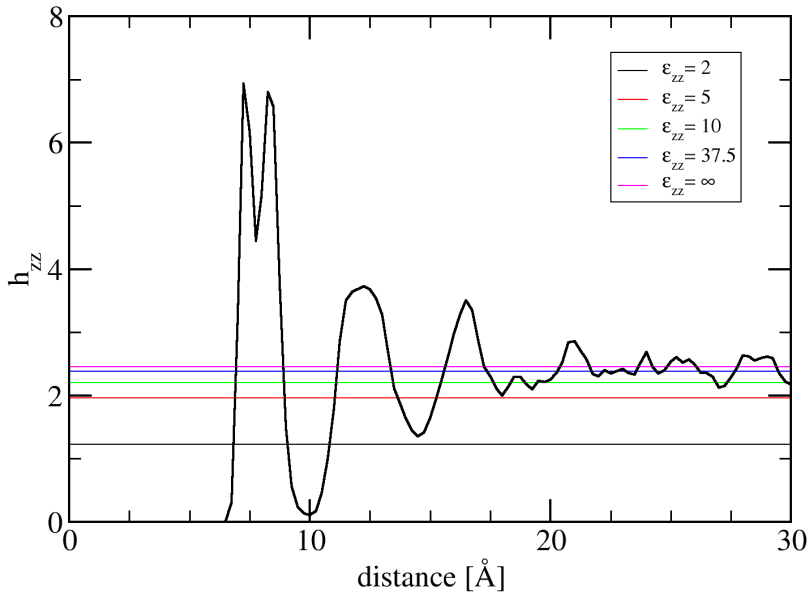

Figure 6.5: Shown is the quantity $h_{zz}(z)$ as defined by Eq. 27 in Ref.[60]. The horizontal lines represent the corresponding dielectric constant. The bulk value matches the expected ($\epsilon = 37.5$) value, but the strong fluctuations of $h_{zz}(z)$ near the interface lead to unphysical values of the local dielectric constant. This suggests that a local linear model of dielectric response is inadequate.[61]

solvent.[47] Here, we observe that the structure of the liquid, as illustrated with a density profile in Fig. 6.6, is indeed reflected in the distributions of the ions. The fact that iodide is preferentially found in the layered solvent close to the interface, while the other ions are rather expelled from that region, must be related to the size of the ion. Indeed, in the model employed, the van der Waals parameters are the only differences between the ions. Iodide, having roughly the size of an acetonitrile molecule, can be better accommodated in the layers than the other, smaller ions, which disrupt the layered solvent structure more. For other solvents, the match between solvent and iodide might be less favorable, and the solvent layering less pronounced. The enhanced iodide concentration at the interface might thus be a feature that is particular to acetonitrile, and might contribute to the efficiency of this particular electrolyte. The effect we observe persists at experimentally rel-

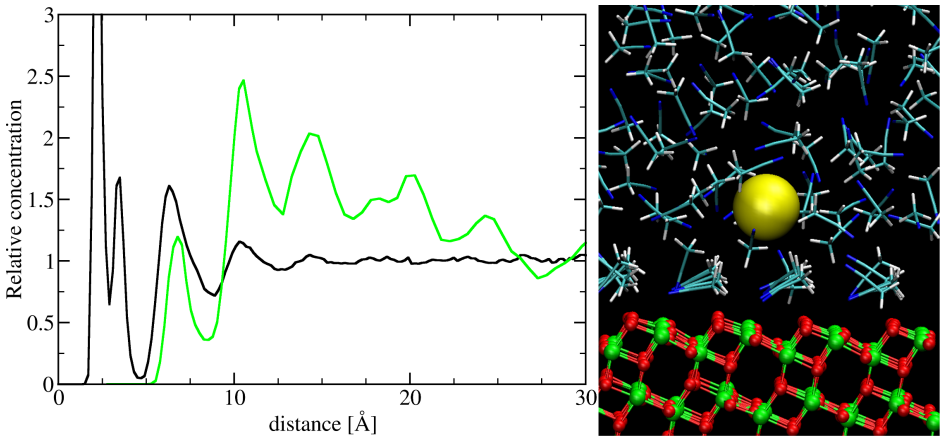

Figure 6.6: The left graph shows the relative concentration profile of iodide compared to the probability density of the acetonitrile central carbon atom. The layering of the solvent (black line) near the interface clearly correlates with the peaks in the ion distribution (green line). The image on the right shows iodide near the interface, approximately at the location of the first maximum in the probability density profile.

evant concentrations of the ions, which we illustrate in Fig. 6.7 with ion profiles obtained by direct MD simulation at a 0.5M concentration of I^- and tetrapropylammonium (TPA^+). The I^- concentration profile obtained in this way agrees very well with the profile obtained from TI. Furthermore, an alternation between layers of high anion and high cation concentration can clearly be observed. Inspection of the molecular dynamics simulations also revealed that diffusion of I^- is facile parallel to the surface, while perpendicular motion is clearly hindered by free energy barriers. Our molecular dynamics data also shows that the first solvent layer, which features strong interactions of the nitrile groups with the oxide, is never penetrated by ions. This suggests that such a solvent layer can effectively passivate the surface, reducing the back electron transfer to solvent species such as I_2^- or I_3^-. A solvent which eliminates these losses improves device efficiency.

6.3 Methods

All density functional theory (DFT) and force field (FF) based molecular dynamics(MD) simulations were performed using the CP2K program package[25]. DFT calculations are based on the hybrid Gaussian and plane

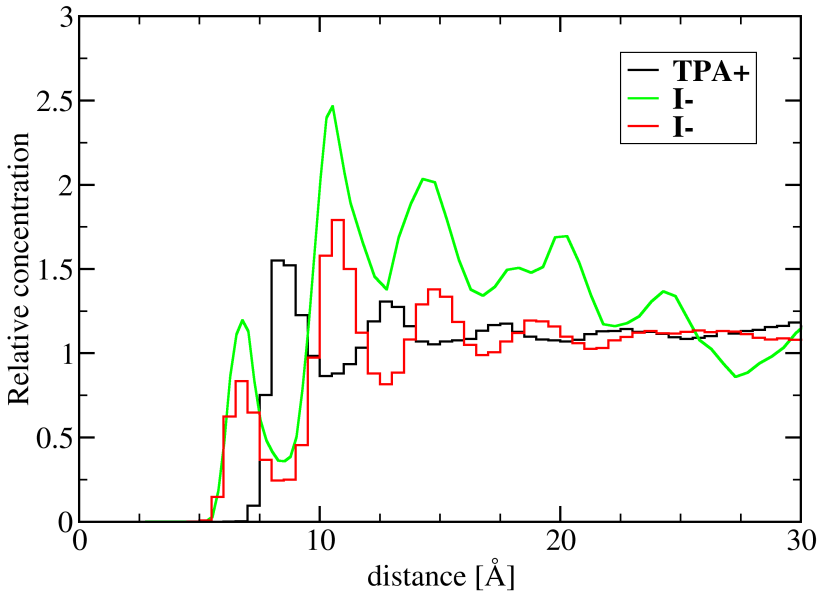

Figure 6.7: The concentration, relative to the bulk value, of iodide and tetrapropylammonium (TPA^+) as a function of the distance to the anatase (101) surface. The binned data is obtained from molecular dynamics simulation at a typical salt concentration (0.5M), while the contiguous line is as in Fig. 6.4

wave (GPW) scheme with the Perdew-Burke-Enzerhof[29] exchange correlation functional and corresponding pseudo potentials[31, 32]. The solvent has been described with Gaussian basis sets of double ζ quality with one polarization function, while a more accurate triple-ζ basis with double polarization has been employed for the dye[26]. The exception are Ruthenium and Iodine, for which highly accurate molecularly optimized [55] basis sets have been used. The plane wave density cutoff was set to 280 Ry. The FF for acetonitrile is based on point charges, van der Waals interactions and harmonic bonded terms as found in literature[38] , TiO_2 was modeled using point charges and a Buckingham potential from literature[37], while the interaction between both species has been described using the van der Waals parameters obtained before. Parameters for the cations have been taken from literature[64, 65], for fluoride we employed $\sigma = 2.90$Å and $\epsilon = 0.105$kJ/mol, for chloride $\sigma = 4.04$Å and $\epsilon = 0.157$kJ/mol, for iodide $\sigma = 4.60$Å and

$\epsilon = 0.438$kJ/mol similar to Ref.[66], the van der Waals interactions between ions and solvent are obtained by the usual combination rules, while only electrostatic interactions are retained between oxide and ions. For classical pre-equilibration of solvent-dye geometries the dye has been parameterized based on the generalized Amber force field. In the first part of this work, constrained ab initio MD is used to calculate the binding free energy of I^- and I_2^- to the oxidized dye solvated in acetonitrile. These simulations have been performed in the microcanonical ensemble(NVE), using a timestep of 0.5fs, with an initial temperature of 300K. The first 5 ps of first principles MD have been considered equilibration. For each value of the reaction coordinate, simulations in the range 5-8 ps have been performed. The initial configurations have been obtained by immersing optimized gas phase structures of the dye in 68–70 molecules of acetonitrile, and equilibrating the solvent molecules only with FF based MD for 1ns in the constant pressure ensemble (NPT). The equilibrated unit cell is approximately 20x20x20$Å^3$. The classical simulations in the second part of this work are NVE, employ a 0.5fs timestep, and are at approximately 300K. For the TI simulation, systems contain approx. 3700 atoms of which one ion (22.7x20.5x113.0$Å^3$), and 800ps of dynamics has been used for averaging the constraint force in each point of the thermodynamic integration (73 points per profile). The explicit simulation of the electrolyte at a finite concentration, starting from an initially random ion pattern, is an average over eight runs, each approx. 3.7ns, systems contain approx. 21700 atoms of which 56 I^- (155.2x40.9x45$Å^3$). Infrared spectra were measured on a Bruker IFS 66/S FT-IR spectrometer equipped with a dedicated attenuated total reflection (ATR)-IR attachment (Optispec) and a liquid nitrogen cooled MCT detector. All spectra were recorded at a resolution of 2 cm^{-1}. The details of the ATR-IR setup is described elsewhere.[67] The film consisting of a titania nanoparticles (Ti-Nanoxide D, Solaronix) was prepared by homogeneously depositing the particles on a Ge internal reflection element which was in advance coated with a 10 nm TiO_2 layer by physical vapor deposition. Subsequently, the film was dried and calcined at 733 K in air for 1 h and transferred to the ATR-IR setup. 0.2 mM of the N719 dye (Solaronix) in ethanol (> 99.9 %, Merck) was slowly admitted until the signals due to dye adsorption were stabilized. The solution in contact with the dye-adsorbed film was first exchanged with acetonitrile (> 99.9 %, Sigma-Aldrich) and then further exchanged with a 0.05 M iodine (> 99.8%, Fluka) solution in acetonitrile to study the interaction of the adsorbed dye with iodine. The reversibility of the dye-iodine interacting complex was further examined by washing the iodine solution with neat acetonitrile. All the measurements were carried out at 313 K.

6.4 Discussion

To summarize our results, we have proposed a complete mechanism for the N3 dye regeneration, which leads via a barrier free pathway from the oxidized dye and I^- to the reduced dye and I_3^-. Key in this process are stable [dye:I] and [dye:I_2] intermediates that dissociate spontaneously upon addition of further I^- or I_2^-. Furthermore, we find that atomistic properties of the electrolyte, both solvent and ion, play an important role in shaping the distribution of ions near the solid/liquid interface. We find that iodide in acetonitrile is a combination that favors dye regeneration, displaying a marked concentration peak in the region of the dye. At the same time, this solvent passivates the surfaces, and reduces back electron transfer, by preventing direct ion–oxide encounters. This insight contributes to the understanding of the most efficient dye sensitized solar cells constructed to date, and is expected to guide rational design of improved devices.

Chapter 7

Computing the Electronic Structure of the Interface

All modeling so far was restricted to structural and energetic aspects. Therefore it is sufficient to use a simplified description of the semiconductor. As soon as the electronic structure plays a more important role, like for level alignment, electronic excitation or electron injection the system requires a more sophisticated treatment. This chapter will discuss at first the effects of the level of theory, solvent and equilibration on the electronic structure. Next, the quality of the model will be illustrated by providing a way to extract electronic excitation spectra from real time propagation and comparing the results to experimental spectra. Finally preliminary results on the electron injection using this setup will be presented.

7.1 Modeling the Electronic Structure

The calculation of electronic properties often requires the use of a more accurate level of electronic structure theory than for strutural properties. The common methods to calculate time-dependent or excited state properties like linear response TDDFT, real time propagation TDDFT and Ehrenfest dynamics have a significant higher computational cost than those for simulating ground state properties. It is therefore important to find good balance between the cost and accuracy has to be found. A quality check of the employed model and level of theory can be obtained comparing computed ground state properties to experimentally known electronic properties. In this section, this strategy will be used to develop a suitable setup to study the electronic properties in DSSC.
One of the most significant problems in modeling DSSC is the description of the semiconductor using density functional theory. Local functionals suffer from significant self interaction errors, when they are applied to such systems. This leads to a large delocalization of the electrons and a underestimation of the band gap of the semiconductor[68]. Several methods exist, which are used to correct for these errors. The most common choice is the use of hybrid

functionals, which correct by adding a given amount of exact (Hartree-Fock) exchange. As pure Hartree-Fock exchange leads to an overestimation of the bandgap, the results obtained by hybrid functionals depends on the fraction of exact exchange mixed in. A second option is adding a Hubbard like U correction to the functionals. The strength of the potential can be adjusted so that the computed band gap matches the experimental one. The disadvantages of this method are that an additional empirical parameter is introduced and mainly the unoccupied orbitals are affected by this correction. In order to find an appropriate tradeoff between accuracy and computational cost, an analysis of the effects of selected functionals, uncorrected PBE, PBE+U and PBE0 as a hybrid functional, on the electronic structure will be presented.
Next to the functionals, the system setup can influence the computed electronic structure. It is a common approach to model the semiconductor only by a small grain of anatase instead of a slab. This approach allows for smaller simulation cells and is therefore computationally less demanding. As the size of these grain is typically chosen very small, the computed electronic structure strongly deviates from the bulk or experimental nanocrystals and is therefore not suitable to describe the electronic processes at the active interface in DSSC. A better description, of the surface, but computationally more demanding, can be obtained using a slab of TiO_2. Using a slab of TiO_2 two dimensions can be treated periodically and therefore resemble better the bulk properties. It is important to note, that even using a slab of anatase, size effects in the nonperiodic direction can remain. Besides the size effects, surface termination, e.q. with protons can influence the electronic structure of the semiconductor. In the previous chapters the orientation of the adsorbed dye on the surface has been determined. As the electronic structure of the dye depends on its binding mode, this knowledge is essential for the modeling of the electronic properties. Furthermore, the solvent, acetonitrile, turned out to be highly structured in the region of the active interface. This structuring in combination with the dipole of acetonitrile can induce an electric field, changing the electronic properties of the components in this region. For these reasons, the different functional will be tested in combination with different system setups to predict the minimal requirements on the structural aspects to model.
Starting with the dye-semiconductor system, PBE turns out to completely close the band gap of the semiconductor. This result is independent of surface protonation and boundary conditions applied. The reason for this wrong result can be found in self interaction errors. About 0.2 electrons of the dye delocalize into the semiconductor. Changing the simulation setup so that the dye and the semiconductor are separated by about 10Å of vacuum the amount of electrons delocalize reduces to 0.06, but the band gap is not opened. A

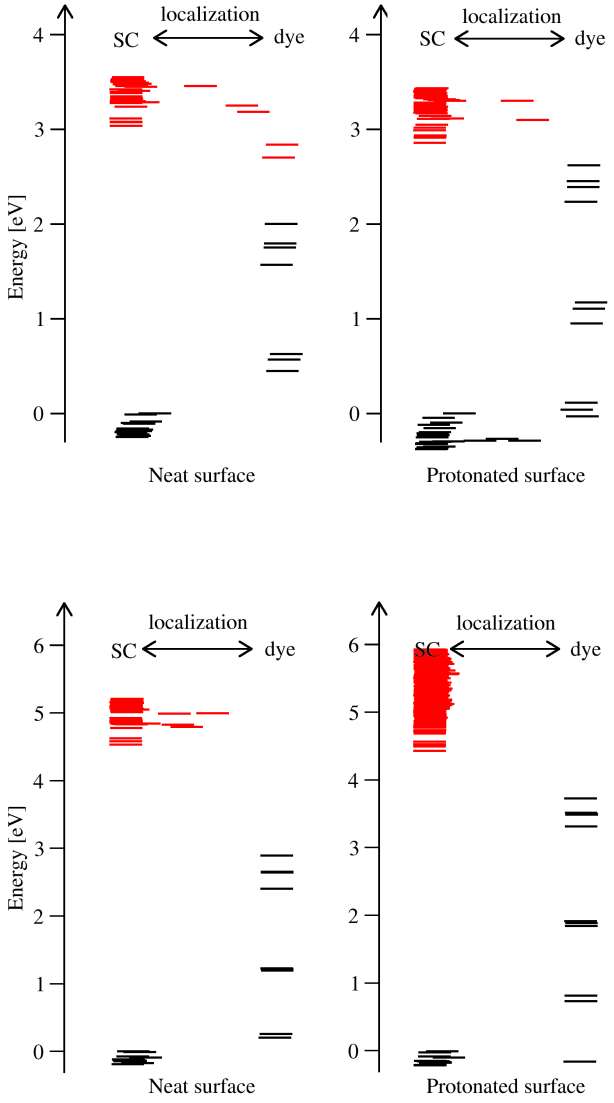

Figure 7.1: Shown are the level diagrams (HOMOs in black, LUMOs in red) for the gas phase dye semiconductor system using the I_2 orientation of the dye. On the top, results using PBE+U for the neat and the protonated surface are presented. On the bottom, the same results for PBE0 are shown. In both cases surface protonation lowers the band gap, and the relative position of the dye HOMOs are shifted upwards.

computationally inexpensive way to open the gap is to apply a +U correction. The penalty potential in these simulations has been applied to the d-orbitals of the titanium atoms, and the magnitude was adjusted to a value of 6.5 eV on a gas phase slab of anatase to match the experimental band gap of 3.2 eV. As it can be seen in Fig 7.1 using the PBE+U functional opens the band gap. The HOMO of this system always corresponds to a dye orbital, which is in agreement with experimental results. Protonating the surface induces a shift of the semiconductor bands to lower energies. As the energy scale in Fig 7.1 is normalized to the valence band edge, the manifests in a shift of the dye levels to higher energies, i.e. the dye-HOMO shifts towards the conduction band edge. Comparing the PBE+U levels with the ones obtained with PBE0 several differences have to be mentioned. The fraction of Hartree-Fock exchange mixed into PBE0 leads to an overestimation of the band gap in this system. Nontheless, as this effect applies to the dye as well, the relative positions of the orbitals are better described using hybrid functionals. Therefore, the level positions obtained with PBE0 will be used as a reference in the following discussion.
Comparing the orbitals localized at the dye, a very similar electronic structure is found for the occupied orbitals according to the same relative position to the HOMO. The major difference is found for the unoccupied orbitals. As the DFT+U correction mainly affects the unoccupied levels of the semiconductor. Hence the first two unoccupied dye states in case of the surface are 0.4 eV below the conduction band edge, while in the PBE0 reference these are already in the conduction band.
For the solvent free system a large effect of surface protonation is found for both PBE+U and PBE0. In both cases the the band gap is lowered by about 0.35 eV. From experiment [69] it is known that surface protonation can affect the band gap by 0.06eV/pH-unit. According to this, the observed change in the calculations would correspond to a change of 6 pH-units. As only two protons are adsorbed on the surface the effect might be predicted to large, which could originate from a missing charge screening due to solvent effects. The different effects of solvation and the importance of high level equilibration will be discussed in the following.
Adding solvent molecules to the simulation induces several important differences. Qualitatively the most important change occurs, when using the PBE functional, which predicts the system in this setup to be semiconducting. As it can be seen in Fig 7.2 the band gap obtained is 2.1eV, which is similar to the value for an isolated slab of anatase with the same settings, but about 1.1eV below the experimental value. PBE+U describes the band gap by construction the best, while PBE0 still overestimates the band gap by 1eV. As these results have been obtained with a protonated surface, the

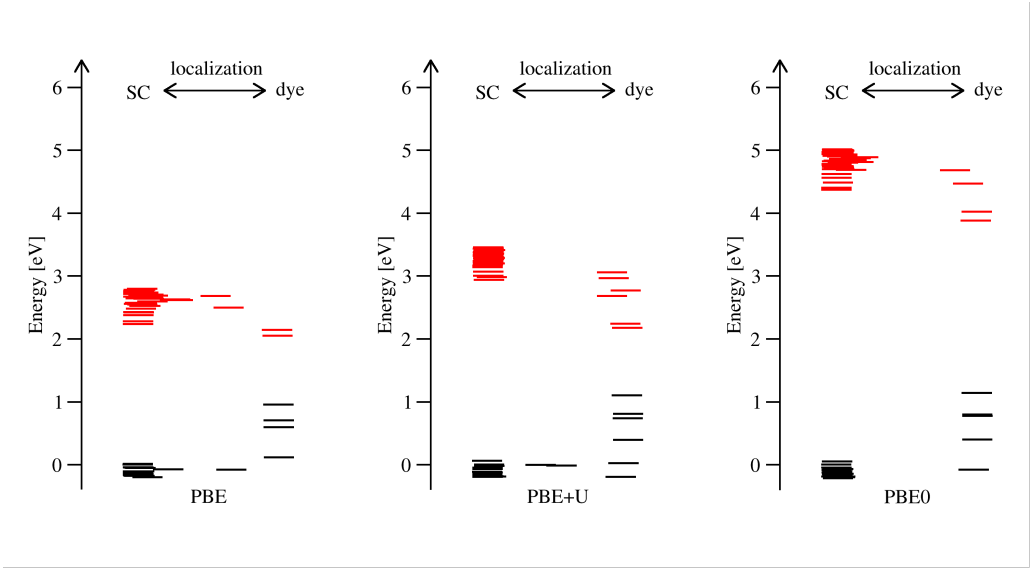

Figure 7.2: Shown are the level diagrams (HOMOs in black, LUMOs in red) obtained with PBE, PBE+U and PBE0 (left to right) for the solvated dye semiconductor system using the I_2 orientation of the dye. Two protons are adsorbed on the surface to obtain a neutral simulation cell. Contrary to the gas phase system, for all three functionals, the first dye LUMOs are located below the conduction band edge.

level diagrams have to be compared to the right sides in Fig. 7.1. Instead of predicting the dyes unoccupied states to be above the conduction band of the semiconductor, they are found about 0.3 eV below in case of PBE and PBE0. The effect is even more pronounced for PBE+U, as the conduction band is pushed further up compared to PBE, therefore, even more, than the two lowest dye unoccupied states appear in the band gap. Partially, this effect can be explained by the structuring of the solvent. As the acetonitrile molecules of the first layer are mostly pointing with their dipole vector towards the surface, an electric potential is created, which shifts the conduction band states to higher energies. Even if this effect can be partially screened by adding additional surface layers, the effect remains for the surface states, and therefore the most important states for electron injection occur at higher energies. Comparing the electronic structure of the orbitals localized at the dye, a very similar result is obtained, with all functionals.
As the configuration for the comparison was obtained from a DFT+U equilibrated trajectory, for which the description of the unoccupied orbitals qualitatively differers, the effect of equilibration using the corresponding function-

als has to be probed as well. In Fig. 7.3 the position and localization of the unoccupied orbitals along a PBE and a PBE0 trajectory is presented. While

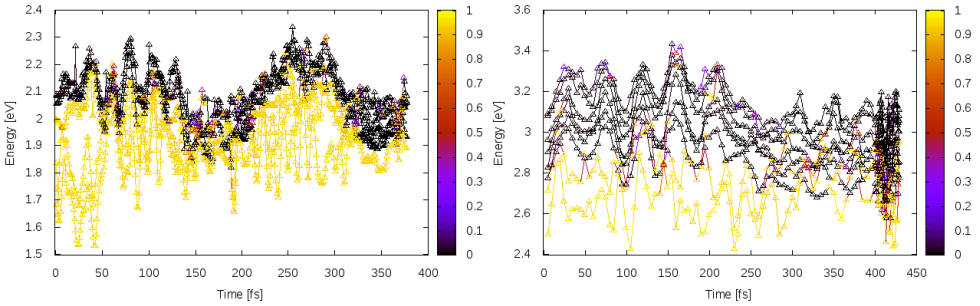

Figure 7.3: Shown is the the level diagram of the unoccupied orbitals, along equilibration using the PBE and PBE0 functional. Starting point of these trajectories have been an equilibrated PBE+U configuration. The color code indicates the degree of localization of the orbital at the dye. In both trajectories, after short equilibration a crossing of surface and dye states can be seen.

the initial state is still the PBE+U configuration, it can be seen, that after 200fs, the dye LUMOs get closer to the conduction band. The PBE trajectory shows strong fluctuations in the level positions leading in a frequency of about 200 fs to a level crossing between dye and surface states. After 350 fs in the PBE0 trajectory, there is a continuous level crossing, for which in the higher resolved part of the trajectory it can be seen to occur on the fs scale. As a 500 fs trajectory is not sufficient for a proper equilibration of this system, it is expected, that the dye levels enter further into the conduction band for longer equilibration times.
From other studies it is known, that the thickness of the slab applied in simulations can affect the band gap of a semiconductor. In order to probe this effect, the band gap of anatase was computed for a model with 3-6 layers using PBE and PBE0. As it can be seen in table 7.1, for less than 5 layers of anatase the bandgap significantly depends on the slab size employed in the model.

number of layers	PBE-bandgap [eV]	PBE0-bandgap [eV]
3	2.53	4.68
4	2.33	4.45
5	2.24	4.35
6	2.19	4.28

For PBE a lowering of the bandgap of about 0.45 eV can be seen. In case of PBE0 this effect is slightly less pronounced with 0.4 eV. For both functionals, the decay is monotonic, and seems to be almost converged for a slab thickness of 6 layers. From these calculations it is not possible to determine the relative shift of valance and conduction band with respect to the dye levels. Nevetheless, comparing the difference in band gap to the level diagrams obtained for the solvated system, a lowering of the conduction band edge of 0.4 eV could bring the dye LUMO from below the band edge in the conduction band. As the increase of the surface size leads to a significantly higher computational cost, this study had to be restricted to a 3 layer model for practical reasons. Nevertheless, this effect can be important for the electron transfer properties, and has to be taken into account, when discussing the further results.

7.2 Excitation Spectra

The computation of excitation spectra of molecules adsorbed on surfaces remains challenging. As it is sensitive to both, functional and the setup of the unit cell, it can be used as a quality check for the description of the electronic structure of the dye. The standard way of computing electronic excitation spectra is linear response TDDFT. A different approach is obtained by propagating the Kohn-Sham equations. Preparing an initial state by exciting the system with a delta pulse

$$v(r,t) = -k_0 p_\nu \delta(t), \tag{7.1}$$

with p_ν the polarization of the pulse, all frequencies of the system are excited with equal weight. Analytically the initial state is obtained by

$$\phi_i(r, t = 0^+) = \hat{T} exp\left(-i \int_0^{0^+} dt \left[\hat{H}_{KS} - k_0 p_\nu \delta(t)\right]\right) \phi_j^0(r). \tag{7.2}$$

Solving this equation, the initial state reads as

$$\phi_i(r, t = 0^+) = exp[i k_0 p_v] \phi_j^0(r). \tag{7.3}$$

Propagating this wavefunction during a finite time, the dynamical polarizability (α_ν) is obtained by

$$\alpha_\nu(\omega) = -\frac{1}{k_0} \int dr^3 p_\nu \delta\rho(r, \omega) \tag{7.4}$$

with $\delta\rho(r,\omega)$ being the Fourier transform of the difference between the time dependent density and the ground state density. The intensity in experimental spectra is then proportional to the average of the imaginary part of the dynamical polarizability over the three spatial directions ν. In case of infinite sampling time, it is valid

$$\delta\rho(r,\omega) = \left\{ \begin{array}{ccc} 0 & \forall & \omega \neq \omega^{ex} \\ \neq 0 & \forall & \omega = \omega^{ex} \end{array} \right\}, \tag{7.5}$$

the excitation energies can also be obtained from any arbitrary initial excitation, but the intensities can not.
In case of a two level system, the wavefunction of the system in an adiabatic basis is given by

$$\Psi(t) = c_\alpha(t)\phi_\alpha + c_\beta(t)\phi_\beta, \tag{7.6}$$

with ϕ_α and ϕ_β as the adiabatic wavefunctions and $c_\alpha(t)$ and $c_\beta(t)$ the time dependent coefficients. The time dependent coefficients c(t) are obtained as

$$c_\alpha(t) = e^{(-i\varepsilon_\alpha t)}c_\alpha \quad and \quad c_\beta(t) = e^{(-i\varepsilon_\beta t)}c_\beta \tag{7.7}$$

with $c_\alpha$and c_β as the coefficients at time zero, by solving the time dependent Schrödinger equation. Thus the analytical form of the time dependent density reads as

$$\rho(r,t) = |c_\alpha|^2\phi_\alpha^2 + |c_\beta|^2\phi_\beta^2 + e^{(-i\Delta\varepsilon t)}c_\alpha^* c_\beta \phi_\alpha^* \phi_\beta + e^{(i\Delta\varepsilon t)}c_\beta^* c_\alpha \phi_\beta^* \phi_\alpha \tag{7.8}$$

with $\Delta\varepsilon = \varepsilon_\beta - \varepsilon_\alpha$. Assuming real coefficients and wavefunctions the latter part of this equation reduces to

$$\rho^{prod}(r,t) = 2c_\alpha c_\beta \phi_\alpha \phi_\beta cos(\Delta\varepsilon t) \tag{7.9}$$

As this is the only time dependent part of the time dependent density, this will be the only contribution to a Fourier transformation. In the discussion above it was shown, that applying an arbitrary excitation to an initial wavefunction, and taking the Fourier transform of the propagated density, only nonzero values are obtained for ω matching an excitation energy. Applying a filter on the Fourier transform at this position in frequency space and transforming the density back to real space, only the time dependent density of the transition with the filtered $\Delta\varepsilon$ remains. It can be shown, that this density is the same as applying the Fourier transformation and backtransformation to a two level system with ϕ_α being the ground state and ϕ_β the excited state wavefunction. Therefore, the result from the filtered density is exactly the product density as defined in equation 7.9 of the ground state and the chosen

excited state of the original system.
In our simulations we used the approach above with the setup of the full interface, to obtain the first excited states using PBE and PBE0. The initial excitation, was prepared by perturbing the first four occupied states of the dye were perturbed by contributions from the corresponding unoccupied states. This initial state was then propagated for about 12.5 fs, resp.

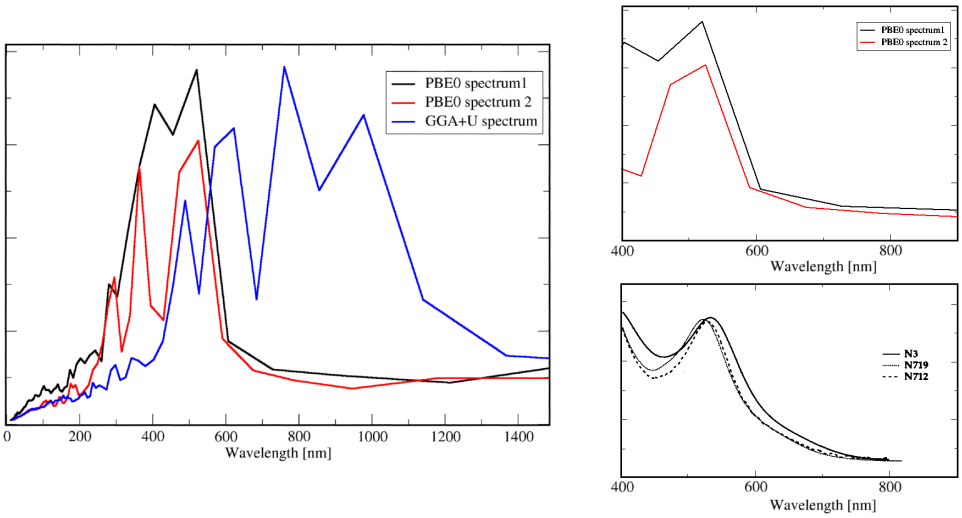

Figure 7.4: Shown is on the left a comparison of the excitation spectra obtained with PBE(blue) and PBE0(black and red) for two different configurations. On the right, the PBE0+U (top) spectra are compared to experimental excitation spectra (bottom)[46]. While PBE+U is significantly shifted to higher wavelength, a good agreement between PBE0 and experiment is found

15.5fs using a time step of 0.2 a.u. for the PBE0 functional. In case of PBE+U, the electron density has been propagated for 20 fs using a time step of 0.125 atomic units. As it can be seen, the spectra obtained with the PBE0 functional are similar for the two different initial configurations. The peak positions for these trajectory are in good agreement with the experimental spectrum of the N3 dye adsorbed on anatase[46]. The PBE+U results are shifted to higher wavelength which originates from the underestimation of the HOMO-LUMO gap of the dye. From these spectra it becomes obvious, that a hybrid functional treatment is necessary to model the electronic properties, as local functional predict the wrong energy scales.
In equation 7.9, the analytical form for the result of the backtransformation

of the filtered Fourier transformed of the density is obtained. Expanding the ground state as a linear combination of the occupied orbitals, and analogue the excited state as linear combination of unoccupied orbitals, a one particle approximation to the ground and excited state wavefunction can be obtained by fitting the density according to 7.9. As an example for this approach, the ground state and the excited state for the filtered signal at 520nm are presented in Fig. 7.4. Here, it can be seen, that the ground state is mainly located at the Ruthenium center and the SCN-ligands, while the excited state is almost fully localized at a single bipyridyl ligand. Experimentally, this excitation is assigned to a metal to ligand charge transfer, which is consistend with the orbital picture shown below.

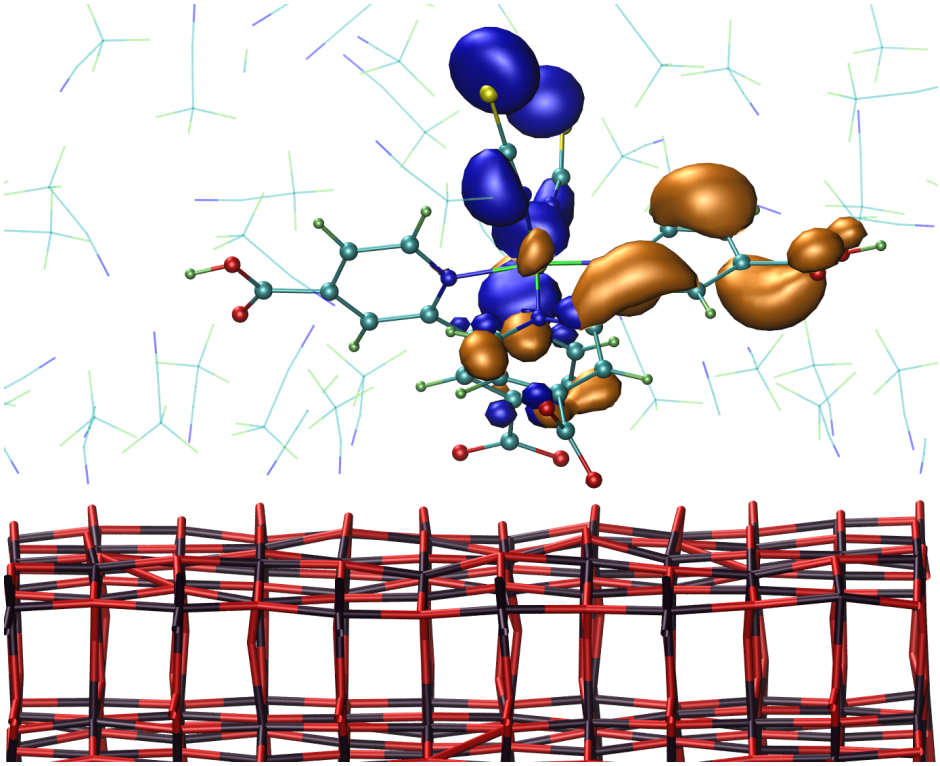

Figure 7.5: Presented is the positive part of the ground (blue) and excited state (orange) wavefunctions obtained by a a fit of the product density defined in Eq. (7.9). The ground state wavefunction is mainly located at the metal center and the SCN-ligand. Contrary, the excited state wavefunction is located at the biisonicotinic acid ligand.

7.3 Electron Dynamics

The electron injection process in DSSC proceeds from the excited state of the dye into the the conduction band of the semiconductor. There are two different pathways through which the electron can enter the surface[70]. One is the adiabatic pathway, which can be described using Marcus theory of electron transfer. The adiabatic electron transfer occurs via an upwards level crossing of the excited state and an unoccupied state. In this case the level occupied by the electron remains at the same relative position, but the nature of the state changes. Hence adiabatic electron transfer can be very fast and efficient, even if only few states are close to the excited state. Next to the adiabatic path, the nonadiabatic path is a possible mechanism for electron transfer in DSSC. Instead of a level crossing, the coupling between the different levels is responsible for the electron transfer. During nonadiabatic electron transfer not only the nature of the state is changed but its relative position as well. The theoretical description of this process is obtained by Fermis Golden Rule. As in this type of electron transfer more than one state is involved, the transfer gets more efficient as more states are involved. For an efficient nonadiabatic electron transfer in DSSC the dye levels have to be significantly above the conduction band edge.
Both processes can be studied using Ehrenfest molecular dynamics (EMD) [71]. Even though it is known, that EMD leads to the wrong equilibrium limits, it is a suitable tool to study the short time evolution of ions and electrons[72]. Here, the evolution of an electron excited from the dye HOMO to the dye LUMO is investigated. In Fig 7.6 integrated excess electron density on the surface along the EMD trajectory is shown, once using the PBE functional and once PBE0. The most obvious feature of both trajectories is the small amount of an electron transferred to the surface. As not only electron injection, but the backreaction is possible as well, this behavior indicates a fast electron backtransfer from the surface to the dye. This is contrary to experiment in which the electron injection is significantly faster than the recombination, but can be understood looking at the level diagrams in Fig 7.2. In the model employed in these calculations, the unoccupied dye orbitals are located around the conduction band edge. Therefore electron relaxation to surface states below the dye LUMO is not possible, and adiabatic electron back transfer becomes favorable. Furthermore, the nonadiabatic electron transfer is expected not to be efficient in this model, as only few surface states can serve as acceptor states. Assuming an equilibrium between back transfer and injection, the amount of electrons at the surface can

be computed as

$$e_S^-(t) = 1 - \frac{k_{bt} + k_{inj}\exp[(k_{inj} + k_{bt})t]}{k_{inj} + k_{bt}}, \tag{7.10}$$

with $e_S^-(t)$ being the amount of electrons transferred at time t, k_{bt} the rate of back transfer and k_{inj} the rate of injection. Applying an exponential fit to the injection curves in Fig. 7.6, an estimate for the injection rate constants can be obtained. For the GGA trajectory a value of about 0.018 fs^{-1} for k_{inj}, and for PBE0 0.009 fs^{-1} is obtained. This corresponds to injection times of >50fs respectively > 100fs. As the conditions in this setup, with a fast upwards and downwards level crossing of the dye LUMO and the conduction band edge, are fairly unfavorable for an efficient electron injection process, the obtained rate presents a upper bound for the injection times. Hence, these results are in reasonable agreement with experiment, predicting injection times of about 30-55fs[73]. Contrary to injection, the adiabatic backtransfer is favored in this model as the dye LUMOs frequently cross the conduction band edge.

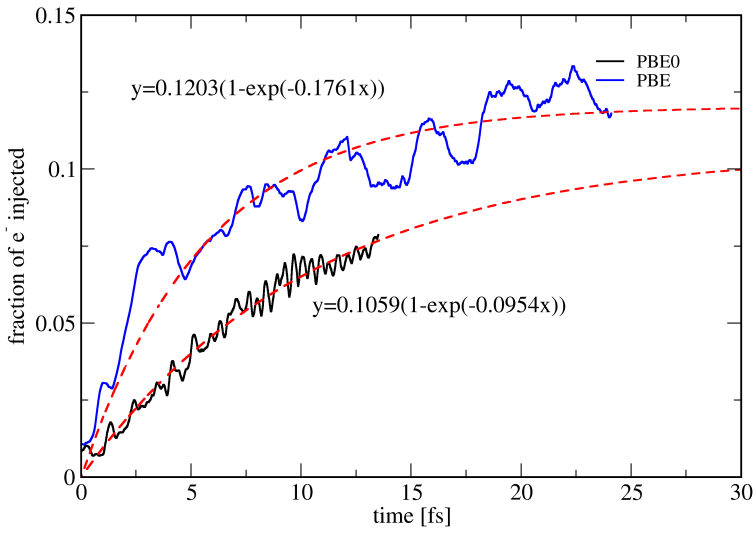

Figure 7.6: Shown is the fraction of an electron transferred to the surface during an Ehrenfest dynamics simulation of the solvated interface using PBE(blue) and PBE0(black). Exponential functions of the form $a_1(1 - \exp(a_2 * x))$ are fitted to these curves to extract the rate constants.

The computed backtransfer time are with 6 fs for PBE and 11 fs for PBE0 orders of magnitude faster than predicted from experiment.

7.4 Methods

All calculations described above have been performed using the CP2K program package. The basis sets and pseudopotentials have been the same for all calculations above. The pseudopotentials and basis sets have been the same as described in the previous chapters. A plane wave density cutoff of 320 Ry was employed for all calculations. Computations on the gas phase dye semiconductor interface were perforemed on the PBE+U optimized I_2-structure obtained in chapter 5. The solvated system was obtained merging the I_2 geometry and the equilibrated MeCN-anatase interface, removing overlapping molecules. This system was than equilibrated for 1ns using force field based molecular dynamics, fixing the positions of the surface and dye atoms. The interactions between MeCN and the semiconductor were modelled using the force field from chapter 4. Standard Amber van der Waals and charge parameters have been used. Solvent-dye van der Waals interactions have than been obtained by applying standard combination rules. The resulting system has been additionally equilibrated for another 10fs using ab initio MD employing the PBE+U functional and the canonical ensemble (NVE). The PBE trajectory presented above was then obtained with a timestep of 0.5fs in the NVE ensemble. In case of the PBE0 trajectory, the multiple time step (MTS) algorithm implemented in CP2K was used. For the reference system on the short time scale PBE forces and a timestep of 0.5 fs has been used. The low frequent PBE0 part has been computed every fifth step to correct the time evolution, which corresponds to an overall time step of 2.5fs. After 400fs of MTS-MD, single timestep MD with the PBE0 functional has been used to validate the correctness of the MTS description.
The configurations for the real time propagation calculations have been taken from the corresponding equilibrated trajectories. The initial perturbation of the wave function was obtained perturbing the first four occupied orbitals of the dye with a 1% contribution of the first four unoccupied orbitals. For both RTP and EMD, the explicite time reversible symmetry propagator and a timestep of 0.125-0.2 atomic units has been used. Furthermore, the matrix exponentials in the propagator has been computed using the adaptive Arnoldi algorithm with an threshhold of 10^{-11}. The initial configurations for the EMD calculations was constructed by choosing states along the trajectories, for which the dye LUMOs are located in the conduction band and an upwards level crossing with the conduction band is expected. The initial

electronic state was prepared by a HOMO to LUMO excitation of a dye electron. All EMD calculations have been performed using the PBE and PBE0 in the local spin density approximation.

7.5 Conclusion

The modeling of the electronic structure of the active interface requires a high level of theory and extended system setups. As it is shown, the exclusion of solvent effects leads to an insufficient screening of charges and therefore to a wrong description of the levels in this system. Furthermore, in a solvent free model, the effect of surface protonation on the band gap of the semiconductor is overestimated, independent of the level of theory applied. Using the DFT+U approach can give, by construction, results for the band gap close to experimental values. Nevertheless, for studies dealing with electronic properties of a system, this approach is not suitable, as the penalty potential only acts on selected orbitals and therefore leads to a misdescription of the electronic structure. The variations in the molecular states found in the Born-Oppenheimer molecular dynamics trajectories indicate, that sampling over many points might be necessary to get reliable results, comparable to experiment. In addition, it is shown, that the use of hybrid functionals is essential to obtain a qualitatively and quantitatively correct description of the excited state properties of the dye.
The timescale for the electron injection process computed with this setup are expected to present an upper bound. For both processes, electron injection and back transfer, the absolute values do not agree with experimental predictions of 30-50 fs for injection[74] and μs scale for recombination[75] . These results can be rationalized, by analyzing the level diagrams in Fig 7.2. At many points of the trajectory, a downwards crossing of the dye LUMO and the conduction band edge is possible. Therefore, electrons reaching the conduction band edge can still be transferred back to the dye. Experimentally, it is found, that the dye LUMO is above the conduction band edge at every time. Hence, electrons at this point are trapped in the conduction band and a significant reduction of the electron back transfer rate can be expected. The 3-layer surface model employed in these calculation led to an overestimation of the gap. As the dye levels remained close to the conduction band edge in the simulations, a lowering of the bandgap might finally shift the levels to the experimental positions and significantly reduce the rate of electron backtransfer in simulations. In addition, as more surface states can interact with the excited state, a stronger nonadiabatic coupling, resp. electron transfer, of the surface and the dye can be favorable.

Chapter 8

Conclusions and Outlook

By a detailed study of the separate interactions occurring at the active interface in DSSC, important insight into processes determining device efficiency could be gained. Although, this study was restricted to a prototypical high efficiency system, many results can be generalized to other device compositions. As shown in chapter two, acetonitrile exposes a strong structuring close to the anatase, which can inhibit the diffusion of the redox couple towards the surface. Therefore, charge recombination reactions between the semiconductor to the redox couple are reduced. Due to the strong interactions and the intrinsic structure already present in the bulk acetonitrile, the passivation of the surface is enforced. Nevertheless, similar effects can be expected in other nitrile containing solvents. The reduced passivation of the surface partially explains the loss in efficiency when using these solvents.
Next to the passivation, the solvent structure is reflected in the concentration of the redox couple close to the surface. This is especially the case for mononuclear ions. While small ions repelled from the surface, the concentration of iodine in acetonitrile is enhanced at 10Å apart from the surface. The enrichment of the reducing agent in this region is favorable effect for a fast and efficient regeneration of the sensitizer after photooxidation. As this concentration profile can be explained by the similar sizes of acetonitrile and iodine, other solvents might not lead to an increased iodide concentration in this region. Nevertheless, other combinations of redox couples and solvents could lead to an enrichment of the reducing agent in a given distance to the surface as well. In combination with the knowledge of the regeneration mechanism of a given sensitizer the device can thus be tuned that the reducing agent is available at high concentration at the active side of the dye. Computational chemistry could give further information by simulating the structural properties of various solvents, respectively ionic liquids and redox couples for the use in DSSC.
The regeneration mechanism proposed in this work includes the formation of several dye-iodine complexes. It was found that this mechanism exclusively works via a sulfur iodine interaction at the thiocyanto ligand for the N3 dye. Comparing the structure of many common inorganic sensitizers, all high performant inorganic dyes contain at least one SCN-group as a ligand.

Therefore, the thiocyanate group is expected to be substantial for inorganic sensitizers and should be kept as a structural feature for new dyes. Furthermore, it is found with this mechanism, that the diiodide radical formed as an intermediate can regenerate the dyes as well, which leads to an additional pathway an therefore to a more efficient regeneration of the dyes.
Combining all results, an almost optimal interplay between the different components in this setup is found. The solvent structure leads to an enrichment of iodide at about 11Å apart from the surface. The orientation of the N3 dye on the surface is such, that the intermediate complex formed during regeneration binds the iodine at exactly this distance. Therefore, all conditions are favorable for a fast regeneration of the dye. These structural aspects are very sensitive to changes in the composition of the cells, and replacement of a single compound can lead to a significant perturbation in these mechanisms. This finding is consistent with experimental studies dealing with different solvents, redox couples or dyes with spacers at the anchoring group. Obviously, the ability of the dye to absorb light and to inject electrons remain the overall limiting step for the performance. Nonetheless, having efficient dyes still requires an efficient regeneration for DSSC to perform well.
It has been shown, that modern computational methods are capable of dealing with complex structures and interfaces. Ground state properties and energetics could be extracted in reasonable agreement with experiment, where reference data were available. Furthermore, the structure of new compounds, like the stable N3-I_2 complex, has been predicted, which have been experimentally detected afterwards.
In the latter part of this work, the focus was on the electronic properties of this system. It was found, that common simplifications of the system like excluding solvent effects, low quality equilibration introduce significant errors. In addition the use of hybrid functionals is essential to obtain a correct description of the electronic structure of the dye. The results obtained for electron injection indicate, that further studies on the electronic properties will have to include a larger surface model than applied in our setup. Moreover, quantitative predictions for the electron injection process require not only a single calculation, but statistics over several different initial configurations. As the setup used in this study already required extensive computer resources, computations on a more realistic setup and extensive sampling in the near future will only become possible through new developments and improved computer hardware.

Bibliography

[1] Nazeeruddin, M. *et al.* Conversion of light to electricity by cis-x2bis(2,2'-bipyridyl-4,4'-dicarboxylate)ruthenium(ii) charge-transfer sensitizers (x = cl-, br-, i-, cn-, and scn-) on nanocrystalline tio2 electrodes. *J. Am. Chem. Soc.* **115**, 6382–6390 (1993).

[2] Tributsch, H. Reaction of excited chlorophyll molecules at eletrodes and in photosynthesis. *Photochem. and Photobio.* **16**, 261–& (1972).

[3] Tributsch, H. Dye sensitization solar cells: a critical assessment of the learning curve. *Coord. Chem. Rev.* **248**, 1511 (2004).

[4] O'Reagan, B. & Grätzel, M. A low-cost, high-efficiency solar cell based on dye-sensitized colloidal tio_2 films. *Nature* **353**, 737 (1991).

[5] Grätzel, M. Photoelectrochemical cells. *Nature* **414**, 338 (2001).

[6] Nazeeruddin, M. *et al.* Engineering of efficient panchromatic sensitizers for nanocrystalline tio2-based solar cells. *J. Am. Chem. Soc.* **123**, 1613–1624 (2001).

[7] Ito, S. *et al.* High-efficiency organic-dye-sensitized solar cells controlled by nanocrystalline-tio2 electrode thickness. *ADVANCED MATERIALS* **18**, 1202+ (2006).

[8] Bergeron, B., Marton, A., Oskam, G. & Meyer, G. Dye-sensitized sno2 electrodes with iodide and pseudohalide redox mediators. *J. Phys. Chem. B* **109**, 937–943 (2005).

[9] Oskam, G., Bergeron, B., Meyer, G. & Searson, P. Pseudohalogens for dye-sensitized tio2 photoelectrochemical cells. *J. Phys. Chem. B* **105**, 6867–6873 (2001).

[10] Sapp, S., Elliott, C., Contado, C., Caramori, S. & Bignozzi, C. Substituted polypyridine complexes of cobalt(ii/iii) as efficient electron-transfer mediators in dye-sensitized solar cells. *J. Am. Chem. Soc.* **124**, 11215–11222 (2002).

[11] Kubo, W., Kitamura, T., Hanabusa, K., Wada, Y. & Yanagida, S. Quasi-solid-state dye-sensitized solar cells using room temperature molten salts and a low molecular weight gelator. *Chem. Comm.* 374–375 (2002).

[12] Wang, P., Zakeeruddin, S., Moser, J., Humphry-Baker, R. & Gratzel, M. A solvent-free, secn-/(secn)(3)(-) based ionic liquid electrolyte for high-efficiency dye-sensitized nanocrystalline solar cells. *J. Am. Chem. Soc.* **126**, 7164–7165 (2004).

[13] Paulsson, H., Hagfeldt, A. & Kloo, L. Molten and solid trialkylsulfonium iodides and their polyiodides as electrolytes in dye-sensitized nanocrystalline solar cells. *J. Phys. Chem. B* **107**, 13665–13670 (2003).

[14] Kohn, W. & Sham, L. Self-consistent equations including exchange and correlation effects. *Phys. Rev* **140**, 1133–& (1965).

[15] Batcho, P. F. & Schlick, T. Special stability advantages of position-verlet over velocity-verlet in multiple-time step integration. *J. Chem. Phys.* **115**, 4019–4029 (2001).

[16] Tuckerman, M. E., Berne, B. J. & Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **97**, 1990–2000 (1992).

[17] Martyna, G. J., Tuckerman, M. E., Tobias, D. J. & Klein, M. L. Explicit reversible integrators for extended system dynamics. *Mol. Phys.* **87**, 1117–1157 (1996).

[18] Runge, E. & Gross, E. Density-functional theory for time-dependent systems. *Phys. Rev.Lett.* **52**, 997–1000 (1984).

[19] Castro, A., Marques, M. & Rubio, A. Propagators for the time-dependent kohn-sham equations. *J. Chem. Phys.* **121**, 3425–3433 (2004).

[20] Kunert, T. & Schmidt, R. Non-adiabatic quantum molecular dynamics: General formalism and case study h-2(+) in strong laser fields. *Eur. Phys. J. D* **25**, 15–24 (2003).

[21] Moler, C. & Van Loan, C. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Rev.* **45**, 3–49 (2003).

[22] Kolafa, J. Time-reversible always stable predictor-corrector method for molecular dynamics of polarizable molecules. *J. Comp. Chem.* **25**, 335–342 (2004).

[23] Reiher, M. & Neugebauer, J. A mode-selective quantum chemical method for tracking molecular vibrations applied to functionalized carbon nanotubes. *J. Chem. Phys.* **118**, 1634–1641 (2003).

[24] Reiher, M. & Neugebauer, J. Convergence characteristics and effciency of mode-tracking calculations on pre-selected molecular vibrations. *PCCP* **6**, 4621–4629 (2004).

[25] The CP2K developers group. http://cp2k.berlios.de/ (2009).

[26] VandeVondele, J. *et al.* Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach. *Comp. Phys. Comm.* **167**, 103 (2005).

[27] Lippert, G., Hutter, J. & Parrinello, M. The gaussian and augmented-plane-wave density functional method for ab initio molecular dynamics simulations. *Theor. Chem. Acc.* **103**, 124 (1999).

[28] VandeVondele, J. & Hutter, J. An efficient orbital transformation method for electronic structure calculations. *J. Chem. Phys.* **118**, 4365 (2003).

[29] Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).

[30] Goedecker, S., Teter, M. & Hutter, J. Separable dual space gaussian pseudopotentials. *Phys. Rev. B* **54**, 1703 (1996).

[31] Hartwigsen, C., Goedecker, S. & Hutter, J. Relativistic separable dual-space gaussian pseudopotentials from h to rn. *Phys. Rev. B* **58**, 3641 (1998).

[32] Krack, M. Pseudopotentials for h to kr optimized for gradient-corrected exchange-correlation functionals. *Theo. Chem. Acc.* **114**, 145 (2005).

[33] VandeVondele, J., Sulpizi, M. & Sprik, M. Electron transfer properties from atomistic simulations and density functional theory. *Chimia* **61**, 155 (2007).

[34] VandeVondele, J. & Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **127**, 114105 (2007).

[35] Lazzeri, M., Vittadini, A. & Selloni, A. Structure and energetics of stochiometric tio2 anatase surfaces. *Phys. Rev. B* **63**, 155409 (1996).

[36] J. Hutter *et al.* `CPMD` (Car–Parrinello Molecular Dynamics): An *Ab Initio* Electronic Structure and Molecular Dynamics Program; IBM Zurich Research Laboratory (1990-2007) and Max–Planck–Institut für Festkörperforschung Stuttgart (1997-2001); see http://www.cpmd.org/.

[37] Bandura, A. V. & Kubicki, J. D. Derivation of force field parameters for tio_2-h_2o systems from a initio calculations. *J. Phys. Chem. B* **107**, 11072 (2003).

[38] Grabuleda, X., Jaime, C. & Kollman, P. A. Molecular dynamics simulation studies of liquid acetonitrile: New six-site model. *J. Comp. Chem.* **21**, 901 (2000).

[39] VandeVondele, J., Lynden-Bell, R., Meijer, E. J. & Sprik, M. Density functional theory study of tetrathiafulvalene and thianthrene in acetonitrile: Structure, dynamics, and redox properties. *J. Phys. Chem. B* **110**, 3614 (2006).

[40] Enjalbert, R. & Galy, J. Ch3cn: X-ray structural investigation of a unique single crystal. b - a phase transition and crystal structure. *Acta Cryst.* **B58**, 1005 (2002).

[41] Marzari, N. & Vanderbilt, D. Maximally localized generalized wannier functions for composite energy bands. *Phys. Rev. B* **56**, 12847 (1997).

[42] Silvestrelli, P. L. & Parrinello, M. Water molecule dipole in the gas and in the liquid phase. *Phys. Rev. Lett.* **82**, 3308 (1999).

[43] Liu, P., Harder, E. & Berne, B. On the calculation of diffusion coefficients in confined fluids and interfaces with an application to the liquidvapor interface of water. *J. Phys. Chem. B* **108**, 6595 (2004).

[44] Kühne, T. D., Krack, M., Mohamed, F. R. & Parrinello, M. Efficient and accurate car-parrinello-like approach to born-oppenheimer molecular dynamics. *Phys. Rev. Lett.* **98**, 066401 (2007).

[45] Finnie, K., Bartlett, J. & Woolfrey, J. Vibrational spectroscopic study of the coordination of (2,2-bipyridyl-4,4-dicarboxylic acid)ruthenium(ii) complexes to the surface of nanocrystalline titania. *Langmuir* **14**, 2744–2749 (1998).

[46] Nazeeruddin, M., Humphry-Baker, R., Liska, P. & Gratzel, M. Investigation of sensitizer adsorption and the influence of protons on current and voltage of a dye-sensitized nanocrystalline tio$_2$ solar cell. *J. Phys. Chem. B* **107**, 8981–8987 (2003).

[47] Schiffmann, F., Hutter, J. & VandeVondele, J. Atomistic simulations of a solid/liquid interface: a combined force field and first principles approach to the structure and dynamics of acetonitrile near an anatase surface. *J. Phys.:Condens. matter* **20** (2008).

[48] Shklover, V., Ovchinnikov, Y., Braginsky, L., Zakeeruddin, S. & Gratzel, M. Structure of organic/inorganic interface in assembled materials comprising molecular components. crystal structure of the sensitizer bis[(4,4-carboxy-2,2-bipyridine)(thiocyanato)]ruthenium(ii). *Chem. Mat.* **10**, 2533–2541 (1998).

[49] Kumar, G., Pan, Y., Smallwood, C. & McAllister, M. Low-barrier hydrogen bonds: Ab initio and dft investigation. *J. Comp. Chem.* **19**, 1345–1352 (1998).

[50] Sasahara, A., Pang, C. & Onishi, H. Stm observation of a ruthenium dye adsorbed on a tio2(110) surafce. *J. Phys. Chem. B* **110**, 4751–4755 (2006).

[51] Fasel, R., Parschau, M. & Ernst, K. Amplification of chirality in two-dimensional enantiomorphous lattices. *Nature* **439**, 449–452 (2006).

[52] Schiffman, F. *et al. Submitted* (2009).

[53] Watkins, M., Trevethan, T., Sushko, M. L. & Shluger, A. L. Designing molecular architecture to control diffusion and adsorption on insulating surfaces. *J. Phys. Chem. C* **112**, 4226–4231 (2008).

[54] Trevethan, T. & Shluger, A. L. Modeling the diffusive motion of large organic molecules on insulating surfaces. *J. Phys. Chem. C* **112**, 19577–19583 (2008).

[55] VandeVondele, J. & Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **127** (2007).

[56] Pelet, S., Moser, J. & Gratzel, M. Cooperative effect of adsorbed cations and iodide on the interception of back electron transfer in the dye sensitization of nanocrystalline tio$_2$. *J. Phys. Chem. B* **104**, 1791–1795 (2000).

[57] Fitzmaurice, D. & Frei, H. Transient near-infrared spectroscopy of visible light sensitized oxidation of i$^-$ at colloidal tio$_2$. *Langmuir* **17**, 1129–1137 (1991).

[58] Clifford, J. N., Palomares, E., Nazeeruddin, M. K., Gratzel, M. & Durrant, J. R. Dye dependent regeneration dynamics in dye sensitized nanocrystalline solar cells: Evidence for the formation of a ruthenium bipyridyl cation/iodide intermediate. *J. Phys. Chem. C* **111**, 6561–6567 (2007).

[59] Privalov, T., Boschloo, G., Hagfeldt, A., Svensson, P. H. & Kloo, L. A study of the interactions between i$^-$/i$_3^-$ redox mediators and organometallic sensitizing dyes in solar cells. *J. Phys. Chem. C* **113**, 783–790 (2009).

[60] Stern, H. & Feller, S. Calculation of the dielectric permittivity profile for a nonuniform system: Application to a lipid bilayer simulation. *J. Chem. Phys.* **118**, 3401–3412 (2003).

[61] Ballenegger, V. & Hansen, J. Dielectric permittivity profiles of confined polar fluids. *J. Chem. Phys.* **122** (2005).

[62] Marcus, R. On theory of electron-transfer reactions .6. unified treatment for homogeneous and electrode reactions. *J. Chem. Phys.* **43**, 679–701 (1965).

[63] Giewekemeyer, K. & Salditt, T. Counterion distribution near a monolayer of variable charge density. *Europhys. Lett.* **79** (2007).

[64] Åqvist, J. Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem.* **94**, 8021–8024 (1990).

[65] Hrobárik, T., Vrbka, L. & Jungwirth, P. Selected biologically relevant ions at the air/water interface: a comparative molecular dynamics study. *Biophysical Chemistry* **124**, 238 – 242 (2006).

[66] Zhang, F. & Lynden-Bell, R. Interactions of triiodide cluster ion with solvents. *Eur. Phys. J. D* **34**, 129–132 (2005).

[67] Urakawa, A., Wirz, R., Burgi, T. & Baiker, A. Atr-ir flow-through cell for concentration modulation excitation spectroscopy: Diffusion experiments and simulations. *J. Phys. Chem. B* **107**, 13061–13068 (2003).

[68] Deskins, N. A. & Dupuis, M. Electron transport via polaron hopping in bulk tio2: A density functional theory characterization. *Phys. Rev. B* **75** (2007).

[69] Redmond, G. & Fitzmaurice, D. Spectroscopic determination of flat-band potentials for polycrystalline tio2 electrodes in nonaqueous solvents. *J. Phys. Chem.* **97**, 1426–1430 (1993).

[70] Prezhdo, O. V., Duncan, W. R. & Prezhdo, V. V. Photoinduced electron dynamics at the chromophore-semiconductor interface: A time-domain ab initio perspective. *Prog Surf. Sci.* **84**, 30–68 (2009).

[71] Duncan, W. R. & Prezhdo, O. V. Theoretical studies of photoinduced electron transfer in dye-sensitized tio2. *Annu. Rev. Phys. Chem.* **58**, 143 (2007).

[72] Parandekar, P. & Tully, J. Detailed balance in ehrenfest mixed quantum-classical dynamics. *J. Chem. Theo. Comp.* **2**, 229–235 (2006).

[73] Benko, G., Myllyperkio, P., Pan, J., Yartsev, A. & Sundstrom, V. Photoinduced electron injection from ru(dcbpy)(2)(ncs)(2) to sno2 and tio2 nanocrystalline films. *J. Am. Chem. Soc.* **125**, 1118–1119 (2003).

[74] Kallioinen, J., Benko, G., Sundstrom, V., Korppi-Tommola, J. & Yartsev, A. Electron transfer from the singlet and triplet excited states of ru(dcbpy)(2)(ncs)(2) into nanocrystalline tio2 thin films. *J. Phys. Chem. B* **106**, 4396–4404 (2002).

[75] Tachibana, Y., Moser, J., Gratzel, M., Klug, D. & Durrant, J. Subpicosecond interfacial charge separation in dye-sensitized nanocrystalline titanium dioxide films. *J. Phys. Chem.* **100**, 20056–20062 (1996).